

Determination of specificity and affinity of the Lactose permease (LacY) protein of *Escherichia coli* through application of molecular dynamics simulation

Master's Degree project in Systems Biology
Two year level, 60 ECTS
Spring term 2018
Report Version 2

Lutimba Stuart
b16stulu@student.his.se

Master Programme in Systems Biology

Supervisor:
Magnus Andersson
magnus.andersson@scilifelab.se

Examiner:
Diana Tilevik
diana.tilevik@his.se
Institution of Bioscience
University of Skövde
BOX 407
541 28 Skövde
Sweden

Abstract

Proteins are essential in all living organisms. They are involved in various critical activities and are also structural components of cells and tissues. Lactose permease a membrane protein has become a prototype for the major facilitator super family and utilises an existing electrochemical proton gradient to shuttle galactoside sugars to the cell. Therefore it exists in two principle states exposing the internal binding site to either side of the membrane. From previous studies it has been suggested that protonation precedes substrate binding but it is still unclear why this has to occur in the event of substrate binding. Therefore this study aimed to bridge this gap and to determine the chemical characteristics of the transport pathway. Molecular dynamics simulation methods and specialised simulation hardware were employed to elucidate the dependency of substrate binding on the protonation nature of Lactose permease. Protein models that differed in their conformation as well as their protonation states were defined from their respective X-ray structures. Targeted molecular dynamics was implemented to drive the substrate to the binding site and umbrella sampling was used to define the free energy of the transport pathway. It was therefore suggested that protonation for sugar binding is due to the switch-like mechanism of Glu325 in the residue-residue interaction (His322 and Glu269) that leads to sugar binding only in the protonated state of LacY. Furthermore, the free energy profile of sugar transport path way was lower only in the protonated state which indicates stability of sugar binding in the protonated state.

Popular scientific summary

Proteins are small organic molecules which perform most of their works in our cells and are required for structure and maintaining the function of organs. Proteins are made up of many smaller units referred to as amino acids that attach to one another in form of long chain that dictate the function of that specific protein. The interaction between the cell and its extracellular environment is mediated by membrane proteins that recognise and transport a variety of small particles such as ions, solutes across the membrane using energy that is obtained from various sources. Lactose permease is a sample to the certain class of proteins, understanding the rearrangement and the structure of this protein is very useful to the industry of drug discovery since it depends on the specific binding of small molecules to their targets that are obviously proteins. This will improve on the existing strategies for drug discoveries and innovations for new drugs. Lactose permease a membrane protein extracted from *Escherichia coli* mediates the accumulation of only lactose sugars into the cell.

Many studies have been carried out on this protein on how it carries out its function. The most recent studies advised that in order for this protein to transport lactose sugars, it needs to have an extra hydrogen ion on the glutamic acid residue but they do not clearly state why this is required. The main aim of the study was to find out why this specific amino acid needs to be having an extra hydrogen ion (protonation) for the sugar to bind to its target protein. If we understand why an extra hydrogen ion is needed in the transport cycle of lactose permease, we then have control over this process that we can regulate and determine the amount of sugars and their types that can enter into and out of the cell. We accomplished this by applying a computer based technique known as molecular dynamics which provides a detailed view on every interaction at an atomic level hence mimicking the natural cellular environment. First, protein copies were defined and these deferred in their rearrangements and also in the presence of extra hydrogen ions in respect to the glutamic acid 325. These models were then inserted into membrane system and then molecular dynamics was applied to them. The interacting molecules were tracked in the process of sugar binding to the protein. An extension to this method was the umbrella sampling methods that determines the amount of the work done by the interacting molecules that is the amino acids in the protein and the lactose sugar.

Results from the simulation are presented in a form of sequential snap shots or frames of pictures that represent the occurring event at that specific time period. It was observed that the reason why protonation is required for substrate binding is due to the contact of specific amino acids that is glutamic acid at the position of 269 in the protein and histidine at the position of 322 both mediated by glutamic acid 325. Among the amino acids in the active site, glutamic acid 269 supports sugar binding. In the protonated state glutamic acid 269 interacts with the sugar but not in the absence of the hydrogen. This is because in the absence of hydrogen ion, glutamic acid 269 interacts with histidine 322 hence hindering sugar binding but in the protonated state, this histidine 322 interacts with glutamic acid 325 hence availing glutamic acid 269 for sugar binding. Furthermore the work done during amino acid sugar-interactions were measured both in the protonated and in the absence of the hydrogen ion. It was observed that there is less work done by the protonated protein system meaning that there is stability unlike in the deprotonated protein system. This generated information is beneficial first to the drug discovery industries as well to the scientific community since it bridges a gap in the understanding of sugar transport in the lactose permease protein.

Abbreviations

E325-: Deprotonated state of the protein by Glutamic acid 325^{H-}

E325H: Protonated state of the protein Glutamic acid 325^{H+}

LacY: Lactose permease (LacY) protein of *Escherichia coli*

TMD: Targeted Molecular Dynamics

PDB ID: Protein Data Database Identification

Table of Contents

Abstract.....	i
Popular scientific summary.....	ii
Abbreviations.....	iii
Table of Contents.....	iv
Introduction.....	1
Function of LacY.....	1
Inward facing conformation.....	2
Proposed mechanism of lactose/H ⁺ symport.....	2
Outward facing conformation.....	3
Molecular dynamics simulations.....	4
Material and methods.....	6
Structure preparation.....	6
Choice of starting structures.....	6
Structural modifications.....	6
Build-up of membrane systems.....	6
Simulation setup.....	7
Targeted molecular dynamics.....	7
Data analysis.....	8
Results.....	10
Equilibration.....	10
Cell dimension analysis.....	10
Maintaining constant temperature and pressure of the systems.....	10
RMS distance.....	10
Hydrogen bond distance.....	11
Sugar binding.....	11
Targeted molecular dynamics.....	11
Interaction hydrogen bond distance analysis.....	12
Pore diameter.....	12
Dihedral angle of phenylalanine 27 residue.....	13
Proposed switch-like mechanism in the transport cycle of LacY.....	15
Discussion.....	16
Equilibration.....	16
Sugar binding.....	17
Mechanism of sugar binding.....	18
Energetics of sugar transport.....	18
Ethical aspects and impact to the Society.....	19
Future perspective.....	20

Acknowledgements.....	21
References	22
Appendices.....	26

Introduction

Proteins are essential in all living organisms [1]. These macromolecules are involved in a range of critically important activities and are also structural components of cells and tissues [2]. The cell and its compartments are separated from their environment by biological membranes. Proteins in the membrane mediate the exchange of substances between cellular environments forming the basis of transmembrane traffic in cellular metabolism [1, 3]. Transmembrane proteins accomplish this transport by structural conformational changes that influence various transport mechanisms. Active transport, unlike passive transport, uses the free energy released from Adenosine Triphosphate hydrolysis (primary active transport) or the free energy stored in an electrochemical proton gradient (secondary active transport) to transport substances across biological membranes [2, 4]. This study focuses on determination of the molecular basis for secondary active transport in Lactose permease (LacY) of *Escherichia coli*, which is a prototype system for the major facilitator super family.

Function of LacY

The diverse major facilitator super protein transporters are single polypeptide secondary carriers [5, 6]. These proteins transport a variety of small molecules in various organisms in response to the electrochemical gradient across biological membranes [4, 7]. LacY is a member of this family and is coded for by the LacY gene, the second structural gene in the Lac operon [5, 7]. LacY utilizes an existing proton gradient to shuttle galactoside sugars across the cell membrane [4, 8]. LacY also transduces energy from the downhill translocation of the sugar to power the uphill transport H^+ in the absence of an electrochemical proton gradient ($\Delta\bar{\mu}_{H^+}$) [4, 5]. To carry out these processes, LacY exhibits specific structural rearrangements leading to intermediate states that bind sugar at one side of the membrane and releases it to the other side of the membrane [9, 10]. The structural rearrangements are driven by subsequent protonation/deprotonation as well as sugar binding/release events [5, 11, 12]. Sugar binding (galactose sugar) is observed to involve an induced fit in LacY [13–15]. In this view, side chains on both the C- and N-terminal helix bundles ligate the galactoside sugar [13, 14, 16]. This leads to a transition to an occluded intermediate state that undergoes an alternate access therefore indicating that the galactose sugar must fully bind for transition into an occluded state [17, 18]. Experimental findings over the past years highlight deprotonation as a rate limiting step for lactose/ H^+ symport in the inward-to-outward transition [9, 19, 20]. In addition, the opening of protonated and sugar-bound LacY to the other side of the membrane is the rate limiting step in the outward-to-inward transition [19].

Key residues involved in the lactose/ H^+ symport mechanism have been described which include Glu126, Asn 272, Arg144, Trp151, Tyr263, Glu269, Arg302, His322 and Glu325 [5, 12, 21]. Findings from various biochemical studies have indicated that the replacement of these key residues may cause blockage of active transport and translocation activity [7, 21, 22]. Glu325 plays a central role in galactoside/ H^+ symport and has a pKa of 10.5, a value that coincides precisely with the variation of the affinity of LacY for galactoside as a function of pH [11, 19, 23]. Several amino acids involved in sugar binding and H^+ translocation have been also suggested to be responsible for the alkaline pKa of Glu325 [23, 24]. This means that Glu325 needs to be protonated to bind the sugar.

Therefore the pKa cannot be assigned to a single amino acid side chain in this group. Glu325 is an essential part of the coupling mechanism in LacY since its protonation or deprotonation determines whether or not galactoside binds effectively [19, 23]. Stabilization of the protonated form of Glu325 is likely due to the hydrophobic microenvironment of the transmembrane domain. [25]. However, it has been suggested that deprotonation of Glu325 involve a decrease in pKa [23]. This has been thought to be caused either by the accessibility of Glu325 to water or by the approximation of Arg302 to Glu325 in the establishment of a salt bridge [16, 26].

LacY is specific in the translocation of sugars with a high specificity directed towards the galactoside sugars, unlike other sugars such as glucose [3, 7]. It has been observed that the specificity is highly related to the C4-OH orientation of the pyranosyl ring which is present in the galactopyranoside moiety of lactose but absent in glucopyranosides like glucose [27]. The conformational changes of LacY in response to affinity and specificity of interacting molecules lead to an alternate access mechanism in the coupled translocation of H⁺ and a galactopyranoside [25, 28].

Although LacY has been extensively studied [5], how sugar affinity and specificity are achieved is presently not known which presents a gap in understanding membrane transport and the coupling mechanism between the proton gradient and transport. Therefore the determination of the molecular features underlying affinity, specificity and coupling mechanism can potentially have wide impact in the structural biology field hence forming the basis of this study. In addition, the entire enterprise of drug discovery rests on selective binding of drug molecules to their targets, which in many cases are membrane proteins [29]. Therefore, understanding of their molecular structure is critical and may boost drug discovery to annihilate diseases.

Inward facing conformation

The historical review of the LacY structure reflects the initial X-ray structures of a stable mutant (PDB ID: 1PV6) and a subsequent wild type structure (PDB ID: 2V8N), which both present the same inward-facing conformation [7, 30]. The basic functional unit for the LacY symporter consist of 12 transmembrane helices ordered into six-helical bundles (the N- and C-terminal) linked by a long cytoplasmic loop [5, 7]. The two transmembrane domains exhibit mirror symmetry and surround a hydrophilic cavity. The sugar binding site is positioned in the centre of the molecule and is located at the apex of the hydrophilic cavity [21, 22, 31]. The cavity also holds the H⁺-binding site. Both of the sites are inaccessible from the tightly sealed periplasmic side and only open to the cytoplasmic side[5]. The tight sealing at the periplasmic cavity is mainly influenced by the closeness that exists between helices II and XI and between helices V and VIII [15]. Residues on helices III, VI and IX are positioned far away from the hydrophilic cavity hence may not directly involve in sugar and H⁺ binding [30].

Proposed mechanism of lactose/H⁺ symport

Understanding the mechanism of substrate and proton translocation gives a detailed insight into how lactose permease carries out its function and also how its structure is directly linked to its role in the accumulation of sugars in the cell of an organism. Therefore it's very important to understand this mechanism for any proposed studies on the function of LacY. It has been proposed that the protonation of LacY in an outward facing conformation from the periplasmic side initiates the reaction cycle [5, 11, 21]. The proton (H⁺) has been suggested to reside in the vicinity of His322 and Glu269. Both of these residues form part of an extensive salt-bridge/H-bond network that defines the H⁺ translocation site in the C-terminal domain [11].

In the protonated LacY, Glu269 and Arg144 are required for sugar binding which disrupts the salt bridge between Arg144 and Glu126 that also initiates the outward-to-inward transition. Glu269 has been identified to provide a link between the H⁺ translocation site and the sugar binding site [11].

The movement of Arg144 influences the relocation of His322 towards Glu325 that acts as a final proton acceptor. The sugar is released to the cytoplasm resulting in conformational changes accompanied by the deprotonation of Glu325 which induces an outward facing conformation [32, 33]. This reverse transition is initiated by the formation of an H-bond between His322 and Glu269, and the reaction cycle is complete [11]. Sugar dissociation leads to the accumulation of sugar against a concentration gradient [34]. The entire reaction steps are described in Figure 1 below.

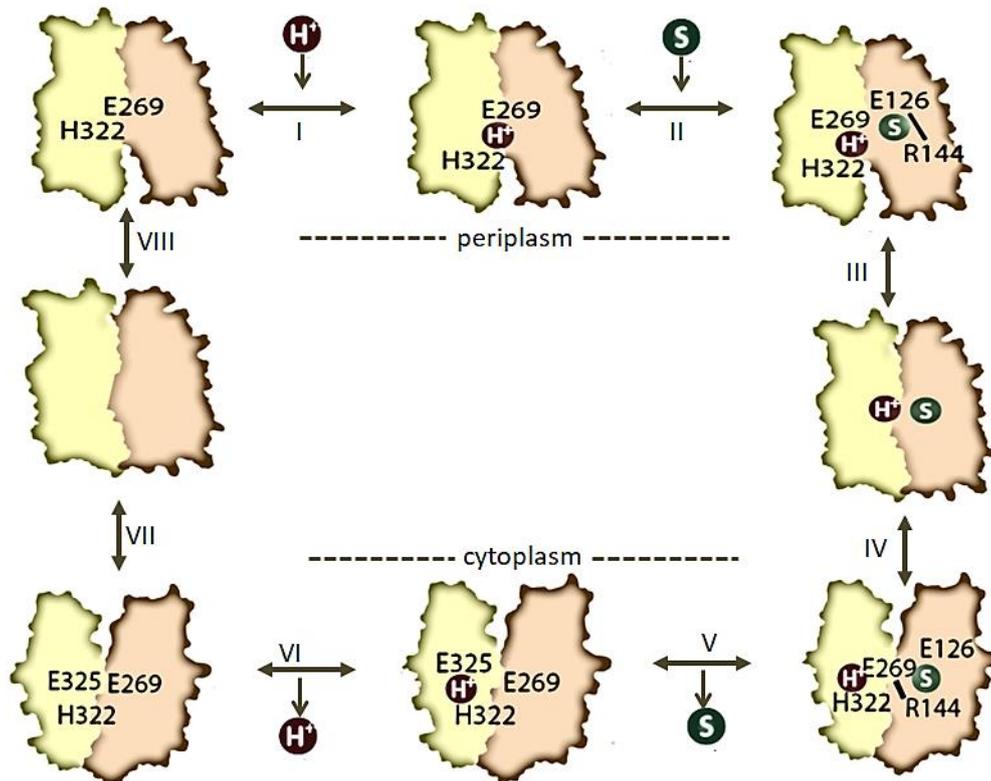


Figure 1. Overview of the proposed model of LacY transport cycle, Steps are numbered sequentially. Step (I) reflects protonation LacY in the outward facing conformation. Step (II) sugar binding and salt bridge interference (Arg144 and Glu126). Step (III) Induced fit occluded conformation following sugar binding. Step (IV) rapid transition of an outward-inward conformation allowing formation of another salt bridge between Arg144 and Glu269. Step V sugar release to the cytoplasm. Step VI deprotonation of Glu325 allowing for an inward-outward transition (VII and VIII). (Drafted from [11])

Outward facing conformation

For many years attempts have been made to trap an outward-facing state of LacY. Recently, a stable double mutant crystal structure was described in an outward occluded state and slightly open to the periplasmic side [9, 10]. This followed the replacement of Gly46 with Trp46 on the N-terminal domain and Gly262 with Trp262 on the C-terminal domain in the periplasmic side. These glycine residues, with similar residues in LacY and among the members of the MSF, are highly conserved [35].

Residues Gly262 and Gly46 in the outward open conformation are significantly important for tighter helix oligomerization allowing close interaction of helices II and VII at the periplasmic side [25]. It has been observed through the experimental body that the replacement of any these residues abolishes dimerization of the transmembrane helices, which traps an outward facing conformation of LacY [32].

The double mutant (G46W/G262W) co-crystallised with high affinity galactopyranosides including β -D-galactopyranosyl-1-thio- β -D-galactopyranoside [PDB ID: 4OAA] and *p*-nitrophenyl- α -D-galactopyranoside (NPG) [PDB ID:4ZYR] provide insight in galactose-specific binding interaction [10]. The interactions of side chains that determine specificity for sugar translocation were observed to be almost indistinguishable in both the sugar analogs [10]. On binding, the sugar goes into an almost occluded but partially open to the periplasmic side with a tightly sealed cytoplasmic side. The incomplete closure was attributed to the bulky tryptophan residues (Trp46 and Trp262), which enable sugar binding without a complete transition into an occluded state [10].

Even more recently, a fully outward-open conformation (PDB ID: 5GXB) was stabilised in an apo-LacY G46W/G262W structure through the use of camelid single-domain nanobodies [36]. Side chains in the binding site that actively determine sugar binding and specificity were observed in the same positions in both sugar-bound and apo states of the protein. Together these three crystal structures rationalize the interactions with LacY for specificity and affinity as discrete properties. However, the underlying molecular details remain unclear. Therefore, determination of the molecular determinants for specificity and of affinity will add to the existing body of knowledge of the molecular basis underlying the coupled transport.

Molecular dynamics simulations

The availability of X-ray structures and computation capabilities with modern algorithms pave way for molecular dynamics (MD) simulation [37]. Such *ab initio* characterization has permitted a detailed understanding of structures and dynamic behaviour of atomistic systems [38]. MD simulations allow tracking of motion of individual atoms on a time scale enabling the elucidation of fast chemical/biological reactions in the transport of sugars and protons across the membrane [39]. MD simulation therefore has complemented the experimentation body by providing a powerful integrative approach in the analysis of biomolecular structures and exploration of conformational dynamics [37].

For example, a previous MD simulation study of a deprotonated LacY state [E325(-)] suggested the closing of the cytoplasmic cavity that destabilises the inward facing state but inducing the outward facing conformation [11]. Therefore deprotonation of Glu325 was suggested to trigger the transition between the inward and outward facing state with more conformational changes at the cytoplasmic side involving helix IV in the N-terminal domain. In addition, an attempt to study the mechanism of permeation was conducted by MD simulation [40]. However, this work was based on the exploration of the periplasmic pathway through the application of steered molecular dynamics in the inward facing conformation (which is closed to the periplasm) and a short simulation time was used for the transport process of lactose sugar.

With access to X-ray structures opened towards the periplasm, we can now for the first time use MD simulation to understand sugar specificity and affinity. In this study, we used the state-of-art MD methods such as targeted MD and umbrella sampling to understand the molecular mechanism and energetic basis of how major facilitator superfamily proteins in biological membranes carry out transport of critical substances. Two LacY systems are established in a POPE lipid membrane in both the protonated (E325/H⁺) and deprotonated state (E325/H⁻). MD simulation were performed using the CHARMM36 [41] parameter set for the biomolecules LacY protein. The LacY membrane systems were composed of lipids, water, ions and a lactose sugar as shown in Figure 2 below.

Together, methods of molecular dynamics simulation and specialised simulation hardware were employed in this study that aimed to determine the molecular elements that govern specificity and affinity as well as proton coupled transport mechanism in Lactose permease system. To accomplish this, the protonation state of Glu325 was explored to find out why Glu 325 must be protonated for substrate binding, therefore protonation of the protein was thought to have an effect on the specificity and affinity of the binding substrate molecule. Understanding the importance of protonation towards the specificity and affinity of the sugar will contribute to the existing body of knowledge of the transport mechanism of LacY. Since LacY is a prototype to the major facilitator super family and shares the same structural rearrangement with majority of members in the family especially the glucose transporter protein, the knowledge in the transport mechanism generated may improve on the existing strategies of drug discovery and innovation of new drugs that are central to human health.

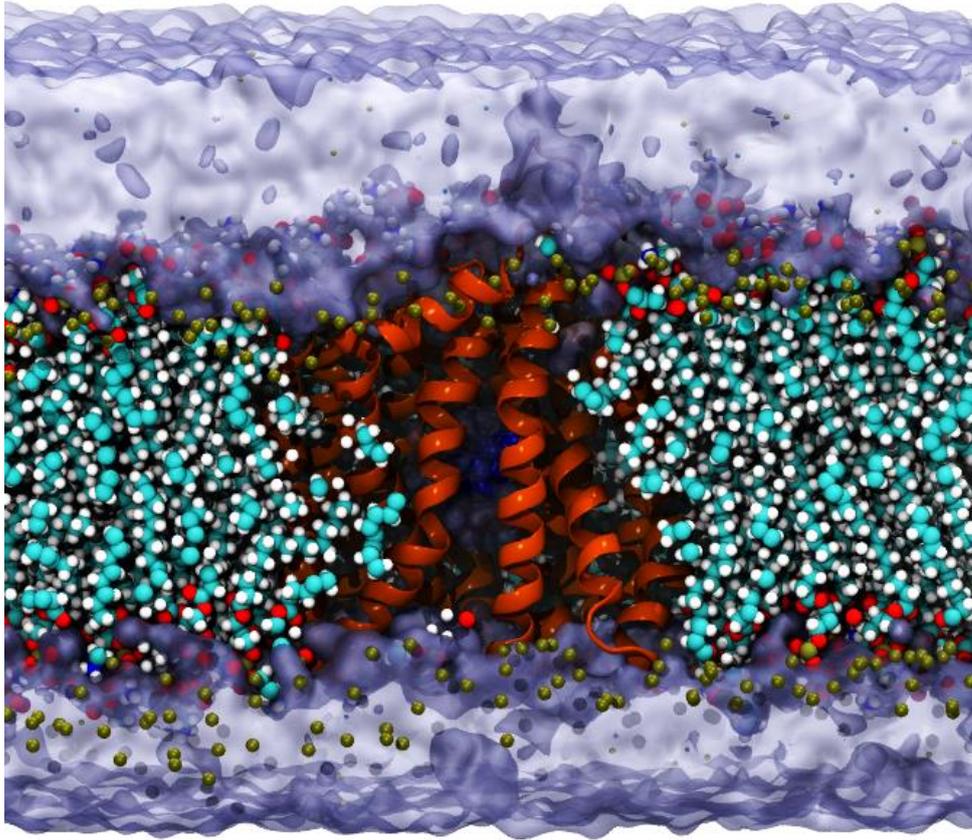


Figure 2. shows the (lactose permease) LacY membrane protein inserted into a phosphatidylethanolamine-phosphatidylglycerol (POPE) bilayer with other system components including water, ions and the sugar molecule. In surface ice blue are the water molecules, phospholipid heads in van der Waals representation presentation coloured in tan, LacY protein presented in orange cartoon, Lipids (POPE) coloured by structure and presented in balls and sticks representation, the lactose sugar in blue balls and sticks buried inside the protein and the ions that dissolved in the water molecules are presented in dots.

Material and methods

Structure preparation

Choice of starting structures

X-ray crystal structures of Lactose permease (LacY) that open towards the periplasmic side were utilised in the determination of specificity and affinity to the selected galactopyranoside sugars. This includes the crystal structure of a conformationally constrained double mutant (G46W/G262W PDBID: 4ZYR) of *E. coli* lactose permease [9]. It was bound to NPG a high affinity sugar molecule. The structure exists as a monomer with a molecular weight of 94.8 kDa and a resolution of 3.312 Å. This state also reveals the sugar binding site describing residues that govern specificity of galactopyranosides. The X-ray structure of LacY-nanobody complex (PDBID: 5GXB) in an outward open conformation was also utilised in this study [36]. This state was stabilized by camelid single domain nanobodies against a double mutant (G46W/G262W PDBID: 4OAA) in the absence of the sugar (apo-state). The structure was determined by X-ray diffraction method with a resolution of 3.3 Å [36]. The protein complex contains chain A as the lactose permease molecule and chain B as the nanobody molecule with a molecular weight of 63.3 kDa. Therefore the X-ray structure with PDBID: 4ZYR and PDBID: 5GXB were used as starting structure for the protein models in the outward facing occluded (periplasmic side almost occluded) state and the outward open (periplasmic side open) state respectively.

Structural modifications

The selected X-ray structures were modified to suit the desired project goals. In both protein structures (PDBID: 4ZYR and PDBID: 5XGB), chain A was selected as the starting coordinate structure. The double tryptophan replacements (G46W/G262W) that were introduced at the periplasmic side in X-ray structures were mutated back to W46G/W262G in the protein models used in the study. The missing loop region that connects the N- and C-terminal at the tightly sealed cytoplasmic side was modelled by satisfying spatial restraints [42]. The modelling program, Modeller 9.19 [50] was used for the production of protein models in both protein states with defined amino acids "(APSSATVANAVGANHS)" in the loop region. Glutamic acid (Glu325) is critical in the proton coupled dynamics and also key in the reaction step of the rapid transition into an outward facing conformation [11]. Therefore protein models that were used in the study differed in the protonation state of Glu325. All histidine residues were defined with protonation state histidine (HSE, proton in the epsilon position).

In the preparation of starting structural files, CHARMM force fields (CHARMM36) [41] were used to reproduce the molecular topology and the selected properties of all atoms in the protein molecule. To accomplish this, VMD (Visual Molecular Dynamics) [43] along with NAMD (Nanoscale Molecular Dynamics) [44] were used for the generation of protein structural files. The "psfgen" plugin, a structural building tool in VMD with topology files (CHARMM 36) were used to define atom types, atom names, bonds and partial charges in the generation of the coordinate files. In addition, coordinates of hydrogen atoms in the protein structures and in water molecules were described using topology files and in protein coordinate files.

Build-up of membrane systems

Definition of the macromolecular environment and incorporation of the protein into lipids to mimic the natural environment is important for molecular simulation. CHARMM-GUI a web based graphical interface for simulation input generation was used to generate the membrane and simulation inputs [41].

A phosphatidylethanolamine-phosphatidylglycerol (POPE) lipid bilayer [45] with dimensions of 120 X 120 X 90 (Å) was generated by CHARMM-GUI. The POPE lipid bilayer was chosen based on the reference to the previous study which elucidated the dependence of LacY dynamics and stability in POPE bilayer environment [11]. The protein models were embedded in the lipids and the entire systems were solvated with the TIP3P water model and ionised with sodium chloride salt at a concentration of 0.15 M. The protein molecule was oriented to the centre of the lipid bilayer. The hydrophobic transmembrane domain in the LacY protein was aligned with the lipid tail locations. The transmembrane region of the protein models was determined using the orientation of protein saver [46] that defined the hydrophobic boundaries with reference to the X-ray structures (PDBID: 4ZYR, PDBID: 5GXB). The segments that exist within the hydrophobic boundaries include 1(9- 31), 2(46- 68), 3(75- 96), 4(105- 130), 5(140- 161), 6(166- 186), 7(222- 244), 8(260- 283), 9(288- 308), 10(313- 337), 11(346- 370) and 12(380- 399) with a total length of 417 residues. All lipid, water molecules and ions within a distance of 3 Å of the protein were removed to avoid steric hindrance during minimisation and equilibration phases of simulation. In this description, four membrane systems were set up that differed in the protonation state of Glu235 (Glu325H and Glu325-) and also in conformation, that is in an almost occluded state (with PDBID: 4ZYR as a starting structure) and an outward open state (with PDBID: 5GXB as the starting structure).

Simulation setup

MD simulations were performed on four LacY membrane systems respectively with reference to their protonation states and their conformation states with NAMD [44], using CHARMM36 [41] force fields. During the initial energy minimisation phase, the systems were simulated with the protein backbone, water molecules and ions constrained for 2 ns following a gradual increase in temperature from 0 to 310 K in a canonical ensemble using a time step of 1.0 fs with the bond length kept rigid. The temperature of the systems was controlled using Langevin dynamics (130 K, with a damping coefficient of 5 ps⁻¹) sufficient to maintain the temperature. Calculations for long range electrostatic force was taken into account by implementing the particle mesh Ewald with a grid size of 140 X 140 X 90 (Å) and a grid spacing of 1.5 Å. Periodic boundary conditions were considered to give the simulation box its shape and size throughout the simulation time thus preventing surface artefacts. The harmonic restraints on the membrane system were released after 2 ns. Another equilibration phase was subsequently conducted for ~ 56 ns timescale until the systems' stability was reached at constant temperature (310 K) and pressure (1 atm). The particle mesh Ewald was also used to calculate electrostatic interactions while short range interactions were cut off at the distance of 12 Å.

A parameterised structure of lactose sugar was introduced into the protein models in the membrane systems at a predefined position. The initial coordinates of the lactose sugar were set at the centre of the periplasmic side of the protein to track sugar pathway towards the binding pocket. Water molecules and ions within 3 Å to the sugar were removed to avoid steric hindrance. Another energy minimization step was performed following insertion of the sugar into the system. This minimisation process also eliminates bad initial contacts. NPT simulation ensemble at 310 K and pressure at 1 atm were subsequently conducted for 5 ns, enough to fully stabilise the systems. During this phase, the sugar, but not the hydrogen atoms, was constrained to avoid undesired movements in the system.

Targeted molecular dynamics

After equilibrating the systems, targeted molecular dynamics (TMD) [44] was utilized to determine the interactions of associated residues and the structural conformations involved in sugar binding along the transport pathway. TMD was performed on the LacY protein model systems, where the sugar molecule (lactose sugar) was guided towards the final target structure by means of steering forces.

The TMD simulation included an additional energy term based on the root mean square deviation (RMSD) of lactose sugar relative to the predetermined targeted structure at finite temperatures. The RMSD measures the interatomic distance between different protein conformational states. Therefore the RMS distance between the current coordinated of the sugar and the target structure was determined at every time step during TMD simulation. The target reference file was essentially a LacY system with bound lactose sugar that was placed in close proximity to binding residues. The residues (Arg126, Glu144, Cys148, Glu269, His322 and Asn272) were located within a distance of 3-4 Å as relates to the sugar binding position of a higher affinity analog NPG.

The fitted and fixed atoms were selected from the protein model while the entire atoms in the sugar were biased. The TMD forces were applied to biased atoms and the RMS distance was linearly decreased from the initial RMSD at the first time step to the final RMSD at the last final step. TMD runs were conducted with a total of 126 TMD atoms, 81 reference atoms and 45 (lactose) biased at a force constant of 7 kcal/mol/Å in the outward facing state and at 7 kcal/mol/Å in the occluded systems at subsequent rates of 1 ns and 10 ns. The simulations were controlled as previously described in the equilibration phase and the resulting trajectories were visualised using the visual molecular dynamics (VMD) software package. The simulations were run using NAMD 2.10 [44] program on a high performance computing system Beskow (CRAY-XC40-MPI) supercomputer at PDC through KTH.

Data analysis

Analyses were done to identify molecular determinants of affinity and specificity as well as the coupling mechanism in LacY transport system. Therefore, the ideal conformational states for sugar binding as well as the importance of Glu325 protonation in the transport pathway of LacY were elucidated. In addition, various analyses were performed to ascertain the stability of the systems throughout the simulation time so as to achieve a local energy minimum in the all simulated systems to prevent structural distortions.

The VMD software package was used to perform visualisation and analysis of output files from NAMD simulations. HOLE program [47] was used in the analysis of the pore dimension that runs through the structural models [47]. Various in-house scripts written in the Tcl/Tk language were also utilised in the analysis of output simulation files.

NAMD simulation program provides algorithms for controlling pressure and temperature of the systems [44]. Therefore, these kinetic properties of the simulation systems were maintained at constant values with reference to the generated statistical ensembles. Corresponding plots were generated to show their distribution and fluctuation over the simulation time. To avoid problems with boundary effects caused by infinite size, periodic boundary condition were considered and the volume of the simulation box was monitored. As the simulation box shrinks when the system is simulated under controlled pressure, the changes of the box were used to monitor the equilibration process. Analysis of the periodic cell parameters that describe the cell dimension was carried out.

The amount by which atoms in the system deviate from the initial coordinates was characterised by the measurement of RMSD for each configuration in the trajectory as a function of simulation time. Thus the RMSD measurement was used for monitoring the integrity and convergence of the simulated system. This also gives insights into the structural conformation throughout the simulation time. Likewise the atoms in the system usually fluctuate around their average positions depending on the temperature factor. The time evolution of the atomic fluctuations for the embedded protein in the systems was monitored for convergence using the RMSF, a useful measure to characterise local changes in the protein.

Hydrogen bonding provides the directional interaction that strengthens the protein structure by conferring rigidity to the intermolecular interactions [48]. Therefore, the time dependent changes in the hydrogen bond lengths of interacting amino acid side chains in the protein were measured and used to monitor the equilibration state of the system. In addition, contacts of hydrogen bonds between the protein and the lactose sugar were defined for to describe directionality and specificity of interactions that are fundamental in the transport pathway. Backbone dihedral angles of side chain residues represent the protein's topology and have been used to compare conformational preferences of a specific residue in transmembrane regions of different states of LacY protein systems. The dihedral angle geometry of phenylalanine 27 (Phe27), a gating residue, was elucidated in order to gain an insight in the sugar transport pathway.

TMD was run with an optimised force constant of 7 kcal/mol/Å that promotes a decrease in the RMS distance of the sugar molecule from the defined state to the targeted state [44]. The RMS distance and the distribution of the applied force constant on the lactose sugar molecule towards its binding site were measured throughout the simulation time by calculating the centre of mass. Water typically participates in sugar recognition and association through hydrogen bonding. Hydration along the sugar pathway to the binding cavity was determined to elucidate protein dynamics during sugar transport. Using the HOLE programme [47] and VMD tools [43], the number of water molecules in the binding site was determined to describe water interactions. TMD lead to generation of the reaction coordinate as the distance from the initial defined sugar position to the targeted binding position[44]. The biased molecular dynamics method, umbrella sampling was used to provide free energy profiles associated with the identified putative sugar entry. Application of this method enables the biomolecular system to evolve from its initial equilibration state therefore enhancing the transitions between different energy minima.

Results

Equilibration

Equilibration in MD is where the Newton's second law is solved for each atom in the system to dictate its trajectory. This allows us to explore the constant energy surface of the system by maintaining the temperature and pressure of the system constant during the molecular simulation to mimic experimental conditions.

Cell dimension analysis

At constant temperature and pressure, stability was achieved by allowing the system to evolve spontaneously for a period of ~ 60 ns enabling the removal of steric hindrances. Atoms in the system encounter different forces during interactions and to counteract this effect, periodic cell boundaries were implemented. The changes in the simulation box were used as a measure to monitor the equilibration of the system.

The relative area (Fig. 1A, Appendix) of the period box in all LacY systems decreases gradually throughout the first 20 ns and levels out there after for the total simulation time of ~ 60 ns. The height (Fig. 1B, Appendix) of the simulation box is observed to decrease initially.

Maintaining constant temperature and pressure of the systems

The LacY systems simulated in the respective protonation and conformation states consisted of varying number of atoms. The outward-facing deprotonated system contained 114,897 atoms and the outward-facing protonated system contained 114,889 atoms. The occluded protonated system contained 111,161 and the occluded deprotonated system contained 111,130 atoms. MD simulations performed at the ensembles (constant temperature and pressure) enabled the generation of several configurations that were used to calculate time dependent properties of the systems. The temperature and pressure plots shown in (Fig. 2, Appendix) were used as a measure to monitor the fluctuations in temperature and pressure throughout the simulation time. The temperature as observed in all systems was maintained within the limits of 310 K. The pressure of the systems was also observed to fluctuate within specific ranges of 1 bar. In the occluded system, fluctuations were observed within 1 Bar and 1 Bar in the out ward-facing system.

RMS distance

The RMSD of the system was measured (Fig. 3, Appendix) by taking the starting structures as reference structures in the calculation. The fluctuation of atoms from their initial coordinate throughout the simulation time represents the convergence and stability of the system. As shown in the (Fig. 3A, Appendix) the protein was constrained initially (2 ns). The RMSD increases rapidly in the first part of the simulation (~ 10 ns) and then stabilises thereafter. The rmsd stabilises around 2.5 Å in the outward-facing E325H and occluded E325- system, around 3 Å in the occluded E325H and around 3.5 Å in the outward-facing E325- system (Fig. 3A, Appendix).

Since the vibrations of atoms are not random but dependent on the structural flexibility [2], the RMSF of each residue was measured for the four systems with consideration of 417 amino acids as shown in (Fig. 3B, Appendix). Convergences were observed in the protein structures in some regions of the protein. In addition, structural deviations were observed among the simulated systems with more frequency in the loop regions especially within the range of 180-220 amino acids. More deviations were observed in the Outward facing E325H (Fig. 6, Appendix).

Hydrogen bond distance

The hydrogen bond distance between specific amino acids in the binding pocket were used as a measure to ascertain the equilibration of the protein. The hydrogen bond distances were considered for Arg144(NH2)-Glu126(OE2) and Arg302(NH2)-Glu269(OE2) as shown in (Fig. 3, Appendix). The inter atomic distance between Arg302(NH2)-Glu269(OE2) increases gradually for the first 5 ns in the (occluded E325-, occluded E325H and outward-facing system) but stabilises thereafter (Fig. 3, Appendix). The outward-facing E325- system shows an initial inter atomic distance decrease for Arg302(NH2)-Glu269(OE2) which stabilises thereafter. There is 1 Å distance difference between the outward-facing state and occluded facing state for the interatomic distance of Arg144(NH2)-Glu126(OE2) (Fig. 3, Appendix).

Sugar binding

Targeted molecular dynamics

The lactose sugar molecule containing 45 atoms was guided to the final target structure by means of steering forces applied. From the TMD runs, the lactose sugars (Fig. 3A) in all the systems were observed to follow a common reaction pathway. This provided reaction coordinates for exploration of the sugar transport pathway which was dependent on each simulated system. Lactose sugar from the initially defined starting coordinates moved linearly along the Z-axis to the binding pocket where it made various contacts with the binding residues. The length of the reaction coordinate varied in every system, with 30.6 Å for the occluded E335H system, 32.1 Å for the occluded E325- system, 29.95 Å for the outward-facing E325- system and 30.94 Å for the outward-facing E325H system.

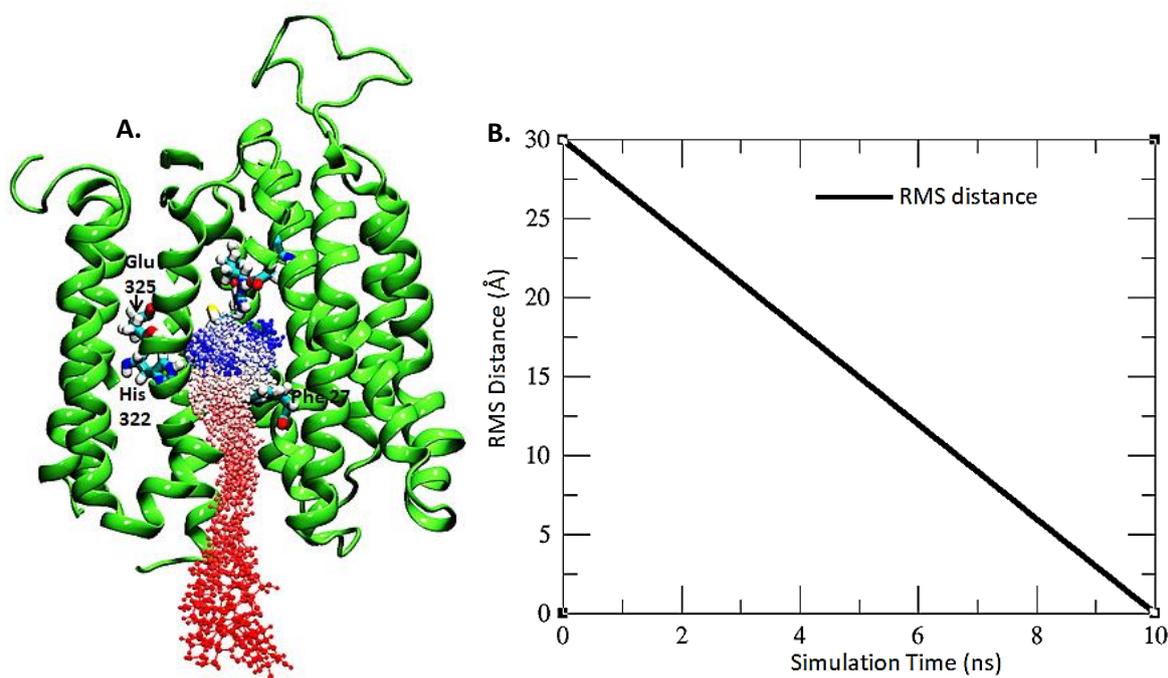


Figure 3. shows the behavior of the lactose sugar following the application of targeted molecular dynamics steering forces at the rate of 10 ns. This presentation is based on the occluded E325- system as similar results were observed in the other systems. Fig. 3A shows the guided movement of lactose (shown in balls and sticks) from the initially defined pose (red balls and sticks) to the binding pocket (blue balls and sticks) throughout the simulation time. Presented in licorice are some of the residues that participate in the sugar translocation pathway. Fig. 3B shows the linear decrease in the rms distance from the initially defined rmsd of 29.95 Å to the final rmsd of 0 Å (targeted binding position of lactose) throughout the simulation time (10 ns).

Interaction hydrogen bond distance analysis

Residues in the water filled binding site of LacY in the model systems such as Glu325, Arg144, His322, Glu269, Arg302, Glu126 and Trp151 were observed to dynamically participate in side chain interactions during H⁺/lactose symport. With reference to the proposed model of LacY transport described in Fig. 1, the salt bridge between Arg302 and Glu325 has been studied in all the model systems for its purpose in sugar binding [11]. The time evolution of Arg302(NH₂)-Glu325(OE₂) distance at the rate of 10 ns described in (Fig. 8, Appendix) shows the inter-atomic interaction between NH₂ group on Arg302 (donor) and OE₂ group on Glu325 (acceptor). In the occluded E325- system (black), the interaction between Arg302 (NH₂) and Glu325 (OE₂) features two possible directions of interactions. After initially restraining the protein structure, the inter-atomic distance is observed to gradually decrease and then stabilises thereafter with a distance of 4 Å. In the occluded E325- system, the distance stabilises at 7 Å for almost 40 ns and far distinct from the E325H system.

In the outward-facing E325H (Fig. 8B, Appendix), the distance stabilises at 8 Å featuring one possible interaction distance that stabilises throughout the simulation time (10 ns). Unlike in the outward-facing E325- system, the inter-atomic distance in the occluded E325- displays two possible directional interactions fluctuating within shorter distances (Fig. 8A, Appendix). In addition, the interaction distance has been observed to coincide with the pK_a of Glu325 (Fig. 8, Appendix) in both the protonated and deprotonated states. In the protonated states, the pK_a is observed to be high as compared to the deprotonated states.

Pore diameter

The structural dynamics and pattern of pore hydration in LacY model systems were measured. A pore diameter of the hydrophilic cavity was therefore computed for the X-ray structures PDBID: 4ZYR (occluded conformation) and PDBID: 5GXB (outward-facing conformation) and for the simulated model systems using the program HOLE [47]. The cavity is observed to differ in dimension, being narrower in the outward facing model systems as compared to the occluded model systems. The variations in pore size in (Fig. 9, Appendix) were monitored throughout the simulation time (10 ns) as an important factor that governs proton transfer events during the transport cycle.

In general, the cytoplasmic region of the hydrophilic cavity remains essentially closed throughout the simulation time (10 ns) in all the simulated model systems. With reference to (Fig. 9, Appendix), the pore diameter profiles indicate that the lumen has two major constrictions both located above (periplasmic region) and below (cytoplasmic region) the binding cavity respectively. In the occluded E325- system (Fig. 9A, Appendix), the pore diameter profile (in pink lines) matches with that of the reference X-ray structure (PDBID: 4ZYR) but not with the X-ray structure (PDBID: 5GXB) and the pore is widest at the binding pocket. A large deviation in pore diameter in the top constriction (periplasmic region) has been observed between simulation profile and that of the reference X-ray structure (PDBID: 5GXB). At the cytoplasmic constriction, the pore diameter profiles in the reference and simulation profile show an agreement. In the occluded E325H system (Fig. 9B, Appendix), the periplasmic constriction and the pore diameter at binding pocket are narrower in comparison to occluded E325- system.

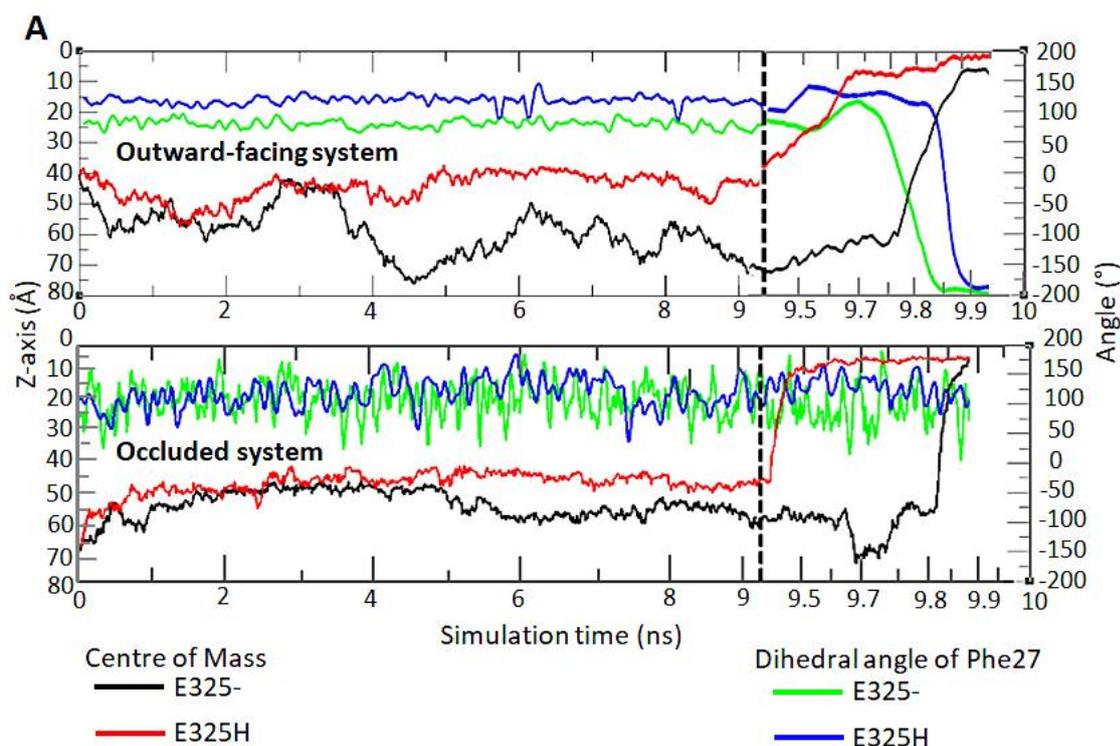
In the outward-facing E325H system (Fig. 9C, Appendix), the cytoplasmic constriction is the wider of the two and the pore is widest at the region around the binding pocket positioned at 5 Å along the Z-axis. The periplasmic constriction matches with the diameter profile of the occluded reference (PDBID: 4zyr) at this particular region but with a large deviation at the pore around the binding pocket. At the binding pocket the simulation diameter profile matches with the reference outward X-ray crystal structure (PDBID: 5GXB) presenting the widest pore diameter.

In the outward E325- system (Fig. 9D, Appendix), the pore diameter of the reference X-ray crystal structure (PDBID: 5GXB) matches with the simulation pore diameter but not with the X-ray crystal structure (PDBID: 4ZYR). At the region around the binding pocket, there is a large deviation in the pore size in comparison with both the X-ray crystal structures that is the PDBID: 4ZYR and PDBID: 5GXB as observed in all the outward model systems. Likewise a large difference in pore size along the Z-axis and at the periplasmic constriction is observed in the reference X-ray structures used. In addition, the periplasmic constriction is largest in the E325H (Fig. 9C, Appendix) in comparison to the E325- (Fig. 9D, Appendix)

Dihedral angle of phenylalanine 27 residue

The LacY model systems present the amino acid phenylalanine 27 (Phe27) that is believed to narrow the periplasmic cavity during the transport cycle [36]. The behaviour of Phe27 has been tracked through the dihedral angle analysis (Fig. 4A). The centre of mass of the sugar representing sugar movement, with Phe27 residue acting as a gating residue has therefore been tracked throughout the simulation time (Fig. 4A).

Calculation of the side chain dihedral angle for Phe27 was in consideration with angle axis CH12: CB-CG. Therefore atom selected for the calculation of the dihedral angle for Phe27 side chain includes CA (368), CB (370), CG (373) and CD1 (374). In the occluded system E325- (Fig. 4A), the dihedral angle for Phe27 is observed to make a slight orientation reassuming the original position hence making way for the sugar to the binding cavity. The Phe27 side chain is observed to open and close after the sugar has passed by this narrow region. This is also observed in the occluded E325- (Fig. 4A) where the side chain paves way for the sugar and then closes back to assume its initial coordinates (Fig. 5, Appendix). Unlike in the outward-facing system E325- (Fig. 4A) the side chain is observed to open at sugar entry to the binding site and doesn't close back to the initial coordinates. This is likewise observed in the outward facing E325H (Fig. 4A). To improve sampling, umbrella sampling was carried out, the reaction coordinate space was divided into windows (Fig. 7, Appendix), and for each independent simulation, the system was allowed to sample within that window for 5 ns. By allowing neighbouring windows to overlap (Fig. 7, Appendix), a PMF was calculated giving statistical weighting to the overlapping regions within windows. Suppl. Fig 7 shows the PMF of the occluded E325- (Fig. 7A, Appendix) and the occluded E325H (Fig. 7B, Appendix), with E325- exhibiting high free energy in contrast to the E325H system.



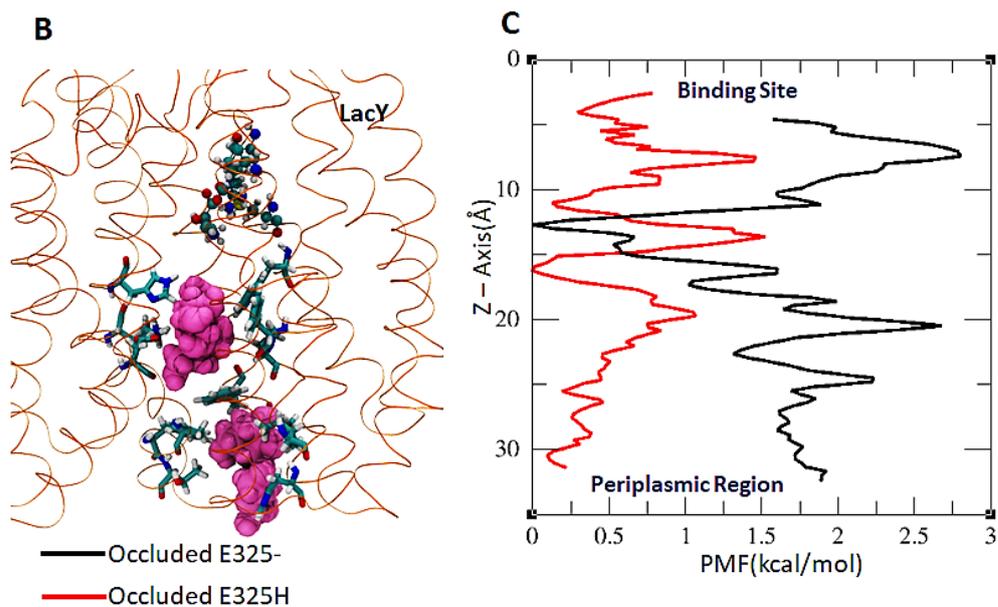


Figure 4. Indicates the centre of mass and the dihedral angle of phe27 residue throughout the simulation time of 10ns, Figure 4A shows the outward-facing system and the occlude system. The centre of mass is represented in red for the protonated systems (E325H) and black for the deprotonated system (E325-). The dihedral angle of phe27 residue is represented in green for the deprotonated (E325-) system and blue for the protonated system (E325H). Fig. 4B shows some of the residues (licorice) causing energy barriers in the transport pathway, in balls and sticks are the residues in the binding site and the sugar is shown in pick *Van der waals* presentation. Fig. 4C represents the energy profile along the Z-axis from the periplasm to the binding site.

Umbrella sampling was used to extract the energy profiles associated with the identified putative sugar entry pathway (Fig.7, Appendix). It's observed that some of the residues (Fig. 4B) along the transport pathway cause an increase in the free energy. The entry pathway has a lower free energy when the protein is protonated (Fig. 4C).

Proposed switch-like mechanism in the transport cycle of LacY

The interatomic distance of specific amino acid in the binding cavity was tracked throughout the simulation time. A residue-residue interaction was observed between specific amino acids in the event of substrate binding.

We therefore present a switch-like mechanism of Glu325 in the residue-residue interaction of His322 and Glu269 in the Fig. 5. In the distance plots in Fig. 5 (A, C and D) below, red represents the protonated system and black represents the deprotonated system. In Fig. 5A, Glu269 approaches the sugar in protonated state (E325H) unlike in the deprotonated state (E325-) of LacY as shown in Fig. 5B with blue representing E325- and red representing E325H. Glu269 in E325- was observed to interact with His322 but not with the sugar as presented in Fig. 5C. In E325H (Fig. 5C), Glu269 does not react with 322 but instead it reacts with Glu325 in Fig. 5D.

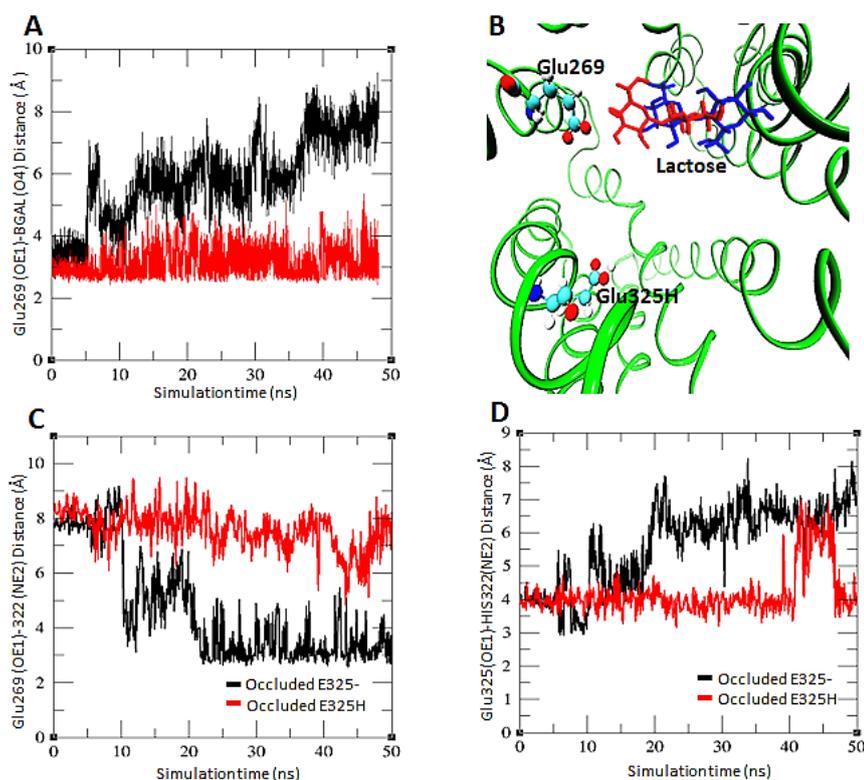


Figure 5. Indicates the residue-residue interaction of His322 and Glu269 in the proposed switch like mechanism of Glu325 in the transport cycle of LacY, Fig: 5A indicated the interatomic distance between the lactose and Glu325, Fig: 5B indicates the proximity of the lactose sugar in both the protonated (in red) and deprotonated (in blue) state of LacY protein (in green). Fig. 5C indicated the inter-atomic distance between Glu269 and His322 and Fig 5D indicates the interatomic distance between Glu325 and His322. In the distance plots red represents the protonated system while blue represents the deprotonated system.

Discussion

Transmembrane proteins mediate the exchange of various substances within the cellular environments of tissues in all life forms[6]. The advancement in structure determination methods has led to the vast number of protein structures that are critical to various cellular processes[49]. This has enabled the elucidation of protein function and to a more extent at an atomic level with the application of the principles of molecular dynamics. Lactose permease a protein model of choice is a prototype to the major facilitator superfamily and share a common fold with majority of members in the family[50]. LacY primary utilizes an existing proton gradient to shuttle galactoside sugars across the cell membrane by exhibiting specific structural rearrangements driven by subsequent protonation/deprotonation as well as sugar binding/release events. The transport mechanism of LacY has been extensively studied through biophysical and biochemical experiments and it has been suggested that it only binds the sugar in only the protonation state[6, 23]. It still remains unclear why this particular protein requires protonation for sugar binding despite the extensive studies carried out[11]. Therefore this study aimed to find the importance of protonation towards the specificity and affinity of the sugar in the transport mechanism of lactose permease protein. To accomplish this, two membrane systems that differed in conformation were modelled and in one case, the protein model was occluded with a narrow opening to the outside.

Simulations of all membrane model systems (Fig. 2) were conducted with CHARMM 36 force fields that enhanced the exploration of large scale LacY conformational changes in the defined simulation time. However, a key limitation in this study was, limited sampling due to the short timescale covered typically of 10 ns[38]. Moreover, limitations in the study were also due to the variations attributed to the force field accuracy (CHARMM 36 force fields)[41].

The structures used differed in their conformation in the way they open towards the periplasmic side. The structure referred to as occluded was more similar to that in the previous studies which was bound a higher affinity analog in the occluded state of LacY[36]. The model protein structure for the outward facing system was more open similar to the structure from the previous studies which was stabilised by a nanobody [10]. In all the defined model system the internal binding site was observed to be positioned in the center of the molecule as also described by previous studies in the determination of substrate binding residues[24].

Equilibration

The protein membrane systems built from the X-ray crystal structures were first equilibrated to obtain stable systems that are free from steric hindrance. The temperature and pressure of the systems was maintained at constant level (Fig. 2, Appendix). This prevented uncoordinated movement of atoms in the system until the equilibrium point was attained. Maintaining the temperature within the limits of 310 K prevented undesired rise in temperature by allowing the extra energy to be released to the surrounding at a physically reasonable rate. Structural deviations from the starting structures were observed in protein models measured by the RMSD based on the C alpha atomic coordinates throughout the simulation time[51]. More fluctuations observed in the first 10 ns (Fig. 3A, Appendix) indicate that the protein structures were searching for conformational stability which was later attained at ~ 60 ns. Deviations observed in all model protein structures were much caused by the flexible loop regions especially the long loop that connects the N- (helix VI) and C- terminal (helix VII)[52]. In addition, a deviation was also observed in the outward facing system caused by a shift in helix V in the E325H model protein hence narrowing the entry path but not in the E325- model protein (Fig. 6, Appendix). This coincides with recent findings where the dependence of protein protonation in the lipid stability was observed to cause a large conformational change at this very helix, there protonation of the protein with respect to Glu325 leads a various conformation changes that contribute to sugar binding[11, 53].

Due to variations in the loop regions, stability of the protein model structures was ascertained through calculation for the hydrogen bond distance key residues in the binding pocket (Arg144, Glu126 and Glu269) and residues that participate in H⁺ translocation (Glu269 and Arg302) (Fig. 4, Appendix). The plot (Fig. 4, Appendix) signifies that the equilibration of the protein model structures was attained throughout the simulation time. This also implies that the internal binding site remains unchanged through time with binding residues remaining in the same positions as defined in crystal structures[25]. It was thus ascertained that all the model systems were equilibrated before the production run as this removed bad contacts from the system which would have caused unnecessary drift in atoms of the model structures.

Sugar binding

Many biological processes rely on the transitions between conformational states of macromolecules.[1] The elucidation of specificity and affinity that is dependent of the protonation nature of LacY in the event of sugar binding involved the application of targeted molecular dynamics which applied a force on the sugar molecule driving it to the targeted binding pocket Fig. 3. TMD was therefore implemented to calculate the reaction paths (Fig. 3A) by continuously decreasing the distance in the starting structure to the targeted conformation with the help of a constraint[54]. It was therefore observed that the gradual decrease trend from the initial RMSD at the first TMD step to the final RMSD at last TMD step was similar in all the systems. With a common sugar path way, it was observed that lactose binding to the pocket is consistent with prior findings from studies of bound NPG and TDG galactopyranoside sugars[9, 10, 36]. This clearly indicates that the transport pathway in the LacY protein is the same irrespective of the conformational state and the protonation state[34]. The galactopyranosyl ring was observed to interact more with residues in the binding cavity with pronounced proximity in the protonated LacY model systems.

The dynamic structural interaction between residues comprising the H⁺ translocation and with reference to the protonation state of Glu325 has been compared between E325H and E325- simulations. It was clearly revealed that Glu325 plays an essential role in galactoside/H⁺ symport and this coincides with the pKa value (Fig. 8, Appendix)[23]. Keeping the sugar in the binding cavity, variations in the pKa value of Glu325 were observed in both the protonated state and the deprotonated state. The inter-atomic distances of Arg302-Glu325 elucidated show various differences along the transport cycle mainly driven by the subsequent deprotonation/protonation events in the bound state of the model systems (Fig. 8, Appendix). It was noticed that stability of Glu325-Arg303 interaction in the protonated systems throughout the simulation time may be influenced by the surrounding residues. It has been suggested before that deprotonation involves a decrease in pKa as this was also observed in the conducted simulations (Fig. 8, Appendix). In the deprotonated system (occluded E325- and outward-facing E325-) a decrease in pKa may thus be attributed to the approximation of Arg302 to Glu325 (Fig. 8, Appendix) as suggested by previous simulation studies in the apo state. Therefore protonation for sugar binding may be dependent on the pKa of Glu325 with deprotonation for the turnover in the transport cycle of LacY.

The structure and pattern of hydration in the pore found at the interface of the N- and C- domain was observed to vary in the LacY model systems (Fig. 9, Appendix). In general the cytoplasmic cavity remains essentially closed in all the simulated systems with deviation in the diameter profile from the crystal structures in the periplasmic region (Fig. 9, Appendix).

The pore diameter is widest at ~ 5 Å along the Z axis as this region constitutes the binding cavity and observed to be consistent with the X-ray structures. This indicates that binding site remains unchanged since it remains open to sugar binding. A significant constriction at 10 Å long the Z- axis presents the gating residue Phe27 as this is in close agreement with the crystal structures of the apo LacY (PDBID: 5GXB) and the sugar bound structure (PDBID: 4ZYR) (Fig. 9, Appendix). Although Phe27

narrows the entrance of the lactose sugar at the periplasmic site, the binding pocket remains freely accessible in all the model systems. Changes in the occluded modelled systems are more pronounced than in the outward-facing state with reference to the narrow region around Phe27. To supplement on the previous studies which suggested that Phe27 narrows the periplasmic region[5], this was observed to be dependent on the conformational state of LacY which are directly influenced by the protonation state of the protein. The outward facing E325H system (Fig. 9A, Appendix) presents a significant decrease of the periplasmic pore diameter that may be attributed to a shift in helix V in response to the protonated Glu325 unlike in the outward-facing E325- system (Fig. 9B, Appendix). In the occluded E325- model system (Fig 9A) there less conformational changes at the periplasmic cavity and thus observed to be in agreement with the X-ray crystal structure unlike in the occluded E325H system (Fig. 8D, Appendix). The significant difference observed at the periplasmic region between the occluded E325H and E325- with more pronounced fluctuations in the occluded E325- may therefore be attributed to the protonated nature of Glu325 side chain.

Mechanism of sugar binding

The coupling between the two states that is the outward-open and inward-open conformation bind the sugar at one side of the membrane and releases it to the other side of the membrane involving the H⁺ translocation [11]. These processes only occur in the protonated state of LacY with Glu325 playing the central role in the H⁺ transfer events [11, 21, 23]. Previous findings suggest that protonation precedes sugar binding though why LacY should be protonated before sugar binding had not been resolved. In Fig. 5, we present a switch-like mechanism of Glu325 in the interaction with Glu269 and H322 that is proposed elucidating need for protonation of LacY in sugar binding. Glu269 that couples between sugar binding and proton translocation is available for sugar binding only in the protonated state (Fig. 5A) unlike in the deprotonated state. This is because a charge on E325 induces a tighter interaction with Tyr236 which makes His322 to move and interact with Glu269 (Fig. 5C) but not with the sugar (Galactoside ring) hence hindering sugar binding in the deprotonated state of LacY. Therefore, the switch-like mechanism of Glu325 in the residue-residue interaction of His322 and Glu269 leads to sugar binding only in the protonated state of LacY.

Energetics of sugar transport

The PMF along the Z -Axis, from the periplasm to the binding site (Fig. 4C) features multiple local energy barriers in the sugar entry pathway. The transition between energy barriers is associated with an increase in the free energy due to sugar-hydrogen interaction as well as hydrophobic contacts during the translocation pathway[9, 55, 56]. The largest free energy barrier was observed in the deprotonated system measuring approximately 2.8 kcal/mol in contrast to the protonated state of LacY (Fig. 4C). Furthermore, the barriers at the periplasmic half channel are smaller reflecting that this channel remains open in contrast to the middle of the channel[17, 21].

Together these findings contribute to the understanding of sugar translocation and H⁺ coupling mechanism in the LacY transport cycle. Affinity was therefore attributed to the galactopyranosyl ring of the lactose sugar since it interacts more in the binding site in the event of sugar binding which was observed to be directly influenced by the protonation of the LacY protein. Therefore with reference to the proposed switch-like mechanism, it was suggested that substrate binding only occurs in the protonated state with specificity influenced by a combination of irreplaceable residues in the binding site which include Glu269, Glu126, Arg302, Arg144, His322 and Glu325. Since LacY is a prototype system for the MFS and exhibits the same basic fold with all other members of the family including the human glucose transporter (GLUT1), these findings may boost drug discovery to annihilate diseases that present the biggest toll on human life hence ensuring the wellbeing of the society.

Ethical aspects and impact to the Society

With reference to the study objective, no ethical concerns and approval were considered since the study was performed through computational simulations (*in silico*). The advantage of conducting *in silico* simulations is that it offers no need for use of animal models, cells and human experiments hence saving time and added costs as well as various regulations. However, existing biochemical and biophysical experimentation methods that form the basis of the determined structures were all under approval by the respective local committees. For example, in the determination of an outward facing conformation of LacY, single-domain nanobodies were used [28, 36]. All practices underlying the development of single-domain nanobodies against a double mutant to stabilise this state were conducted after approval by an ethical committee (Ethical committee at Vrije Universiteit Brussel, VUB project 13-601-1) and also following the guidelines of EU animal welfare [36].

Following previous studies on the structure of Lactose permease from *Escherichia coli*, the starting structures used for the aim of the study were prepared based on their corresponding X-ray structures PDB ID: 5GXB [28] and PDB ID: 4ZYR [9]. These were obtained from the public repository Protein Data Bank (PDB) www.rcsb.org [57]. Therefore all materials used from the database archive were free from all copyright restrictions and was fully accessible for academic and research purpose. The determination of affinity and specificity in LacY through the application of molecular dynamics will contribute to the existing body of knowledge and understanding of the coupled sugar/H⁺ transport mechanism. In addition, this study with the employment of LacY a model system, could give insights which may be applicable in drug design hence benefiting the scientific community. It is also intended that results from this study are to be compared to the already existing biochemical data and thus there no ethical aspects to consider. These findings may boost drug discovery to annihilate diseases that present the biggest toll on human life hence ensuring the wellbeing of the society.

Future perspective

Biomolecules typically consist of a vast number of interacting particles and their properties give a holistic view of the dynamic evolution of the system[7, 39]. MD as a method of choice was utilised to understand the dynamics in sugar binding with reference to the protonated state of LacY. A detailed atomistic view of the molecular elements that govern specificity and affinity in the event sugar binding was obtained.

This study provided a proposed switch-like mechanism of Glu325 that defines why LacY must be protonated for sugar binding to occur hence influences the affinity and specificity of the protein towards substrate binding. Therefore, from now we know that the inter-atomic interactions of Glu269 and H322 influenced by Glu325H determine sugar binding hence bridging a gap in the LacY transport mechanism. With reference to the energetics of sugar transport[56], lower free energy was observed in the protonated sugar pathway unlike in deprotonated pathway which was brought about by the switch-like mechanism of Glu325.

With reference to the short timescale used, future simulations should be based on more numbers of simulation time scales to capture more dynamics in the transport cycle[40]. However, our simulations do provide an atomistic detailed picture of the molecular events relevant to the transport mechanism[21, 34]. Future experimental studies should be performed on these systems to validate the proposed switch-like mechanism in the transport cycle of LacY.

Acknowledgements

First and foremost, I would like to express my great appreciation to the Swedish Institute that awarded me a scholarship within the Swedish Institute Study Scholarship for studies at the Master programme in Molecular Biotechnology at the University of Skövde.

I express my deep gratitude and appreciation to **Magnus Andersson**, Associate Professor, my main supervisor for the guidance, encouragement and advice he has provided through my time as his student. Thank you Magnus for making my career dreams come true by accepting me to be part of your Lab. I have not just learned about scientific research using advanced computational methods but also other aspects of scientific life such writing, you opened up my mind to various career paths, it has been a pleasure and I have enjoyed working with you. I am really grateful for everything; I wouldn't have attained success without your insightful suggestions during planning and development of this research thesis. Thanks a lot Magnus.

I express my very great appreciation to **Diana Tilevik** Ph.D. my examiner for her valuable and constructive suggestion towards my research thesis project. Her willingness to give her time so generously has been very much appreciated. I add that I have been extremely very lucky to have supervisors who cared so much about my work, and who responded to my questions and queries so promptly.

I would like to thank **Erik Lindahl's group (TCBLAB)** at Science for Life Laboratory for all the support provided and also creating a stimulating learning environment. I always learnt more from the journal club meetings and presentations conducted in the group.

I send a vote of thanks to Maria Algerin, Ph.D. for the insightful introduction to the principles of molecular biotechnology and for a warm welcome to the University of Skövde.

Finally, I would like to express my gratitude my parents the **Late Kyeswa Charles and Ms Nakawuka Marion** with all members of my entire family (**Kyeswa family**) for their support and encouragement during my studies and during this thesis. My accomplishments would never have been possible without them. Please be blessed.

Thanks to you all.

References

1. Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., Walter, P.: *Molecular Biology of the Cell* 6e. (2014)
2. Lodish, H., Berk, A., Zipursky, S.L., Matsudaira, P., Baltimore, D., Darnell, J.: *Molecular Cell Biology*. 4th edition. (2000)
3. Cosson, P., Perrin, J., Bonifacino, J.S.: *Anchors aweigh: Protein localization and transport mediated by transmembrane domains*, (2013)
4. Robertson, D.E., Kaczorowski, G.J., Garcia, M.L., Kaback, H.R.: Active transport in membrane vesicles from *Escherichia coli*: the electrochemical proton gradient alters the distribution of the lac carrier between two different kinetic states. *Biochemistry*. 19, 5692–5702 (1980)
5. Guan, L., Kaback, H.R.: LESSONS FROM LACTOSE PERMEASE. *Annu. Rev. Biophys. Biomol. Struct.* 35, 67–91 (2006). doi:10.1146/annurev.biophys.35.040405.102005
6. Quistgaard, E.M., Löw, C., Guettou, F., Nordlund, P.: Understanding transport by the major facilitator superfamily (MFS): structures pave the way. *Nat. Rev. Mol. Cell Biol.* 17, 123–132 (2016). doi:10.1038/nrm.2015.25
7. Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., Iwata, S.: Structure and mechanism of the lactose permease of *Escherichia coli*. *Science*. 301, 610–615 (2003). doi:10.1126/science.1088196
8. Nie, Y., Ermolova, N., Kaback, H.R.: Site-directed Alkylation of LacY: Effect of the Proton Electrochemical Gradient. *J. Mol. Biol.* 374, 356–364 (2007). doi:10.1016/j.jmb.2007.09.006
9. Kumar, H., Finer-Moore, J.S., Kaback, H.R., Stroud, R.M.: Structure of LacY with an α -substituted galactoside: Connecting the binding site to the protonation site. *Proc. Natl. Acad. Sci.* 112, 9004–9009 (2015). doi:10.1073/pnas.1509854112
10. Kumar, H., Kasho, V., Smirnova, I., Finer-Moore, J.S., Kaback, H.R., Stroud, R.M.: Structure of sugar-bound LacY. *Proc. Natl. Acad. Sci.* 111, 1784–1788 (2014). doi:10.1073/pnas.1324141111
11. Andersson, M., Bondar, A.N., Freitas, J.A., Tobias, D.J., Kaback, H.R., White, S.H.: Proton-coupled dynamics in lactose permease. *Structure*. 20, 1893–1904 (2012). doi:10.1016/j.str.2012.08.021
12. Smirnova, I.N., Kasho, V., Kaback, H.R.: Protonation and sugar binding to LacY. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8896–8901 (2008). doi:10.1073/pnas.0803577105
13. Frillingos, S., Kaback, H.R.: The role of helix VIII in the lactose permease of *Escherichia coli*: II. Site-directed sulfhydryl modification. *Protein Sci.* 6, 438–43 (1997). doi:10.1002/pro.5560060221
14. Mirza, O., Guan, L., Verner, G., Iwata, S., Kaback, H.R.: Structural evidence for induced fit and a mechanism for sugar/H⁺ symport in LacY. *EMBO J.* 25, 1177 LP-1183 (2006)
15. Liu, Z., Madej, M.G., Kaback, H.R.: Helix Dynamics in LacY: Helices II and IV. *J. Mol. Biol.* 396, 617–626 (2010). doi:10.1016/j.jmb.2009.12.044
16. Frillingos, S., Gonzalez, A., Kaback, H.R.: Cysteine-scanning mutagenesis of helix IV and the

- adjoining loops in the lactose permease of *Escherichia coli*: Glu126 and Arg144 are essential. *Biochemistry*. 36, 14284–14290 (1997). doi:10.1021/bi972314d
17. Jiang, X., Guan, L., Zhou, Y., Hong, W.-X., Zhang, Q., Kaback, H.R.: Evidence for an intermediate conformational state of LacY. *Proc. Natl. Acad. Sci. U. S. A.* 109, E698-704 (2012). doi:10.1073/pnas.1201107109
 18. Kaczorowski, G.J., Robertson, D.E., Kaback, H.R.: Mechanism of Lactose Translocation in Membrane Vesicles from *Escherichia coli*. 2. Effect of Imposed $\Delta\Psi$, ΔpH , and $\Delta\mu\text{H}^+$. *Biochemistry*. 18, 3697–3704 (1979). doi:10.1021/bi00584a010
 19. Smirnova, I., Kasho, V., Sugihara, J., Choe, J.Y., Kaback, H.R.: Residues in the H^+ translocation site define the pKa for sugar binding to LacY. *Biochemistry*. 48, 8852–8860 (2009). doi:10.1021/bi9011918
 20. Majumdar, D.S., Smirnova, I., Kasho, V., Nir, E., Kong, X., Weiss, S., Kaback, H.R.: Single-molecule FRET reveals sugar-induced conformational dynamics in LacY. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12640–5 (2007). doi:10.1073/pnas.0700969104
 21. Jiang, X., Villafuerte, M.K.R., Andersson, M., White, S.H., Kaback, H.R.: Galactoside-binding site in LacY. *Biochemistry*. 53, 1536–1543 (2014). doi:10.1021/bi401716z
 22. Yin, Y., Jensen, M.Ø., Tajkhorshid, E., Schulten, K.: Sugar binding and protein conformational changes in lactose permease. *Biophys. J.* 91, 3972–85 (2006). doi:10.1529/biophysj.106.085993
 23. Grytsyk, N., Sugihara, J., Kaback, H.R., Hellwig, P.: pKa of Glu325 in LacY. *Proc. Natl. Acad. Sci. U. S. A.* 114, 1530–1535 (2017). doi:10.1073/pnas.1621431114
 24. Vadyvaloo, V., Smirnova, I.N., Kasho, V.N., Kaback, H.R.: Conservation of residues involved in sugar/ H^+ symport by the sucrose permease of *Escherichia coli* relative to lactose permease. *J. Mol. Biol.* 358, 1051–1059 (2006). doi:10.1016/j.jmb.2006.02.050
 25. Zhou, Y., Jiang, X., Kaback, H.R.: Role of the irreplaceable residues in the LacY alternating access mechanism. *Proc. Natl. Acad. Sci.* 109, 12438–12442 (2012). doi:10.1073/pnas.1210684109
 26. Sahin-Tóth, M., Le Coutre, J., Kharabi, D., Le Maire, G., Lee, J.C., Kaback, H.R.: Characterization of Glu126 and Arg144, two residues that are indispensable for substrate binding in the lactose permease of *Escherichia coli*. *Biochemistry*. 38, 813–819 (1999). doi:10.1021/bi982200h
 27. Sahin-Tóth, M., Lawrence, M.C., Nishio, T., Kaback, H.R.: The C-4 hydroxyl group of galactopyranosides is the major determinant for ligand recognition by the lactose permease of *Escherichia coli*. *Biochemistry*. 40, 13015–13019 (2001). doi:10.1021/bi011233l
 28. Smirnova, I., Kasho, V., Jiang, X., Pardon, E., Steyaert, J., Kaback, H.R.: Outward-facing conformers of LacY stabilized by nanobodies. *Proc. Natl. Acad. Sci.* 111, 18548–18553 (2014). doi:10.1073/pnas.1422265112
 29. Overington, J.P., Al-Lazikani, B., Hopkins, A.L.: How many drug targets are there? *Nat. Rev. Drug Discov.* 5, 993–6 (2006). doi:10.1038/nrd2199
 30. Guan, L., Mirza, O., Verner, G., Iwata, S., Kaback, H.R.: Structural determination of wild-type lactose permease. *Proc. Natl. Acad. Sci. U. S. A.* 104, 15294–8 (2007).

doi:10.1073/pnas.0707688104

31. Sahin-Tóth, M., Gunawan, P., Lawrence, M.C., Toyokuni, T., Kaback, H.R.: Binding of hydrophobic D-galactopyranosides to the lactose permease of *Escherichia coli*. *Biochemistry*. 41, 13039–13045 (2002). doi:10.1021/bi0203076
32. Smirnova, I., Kasho, V., Sugihara, J., Kaback, H.R.: Trp replacements for tightly interacting Gly-Gly pairs in LacY stabilize an outward-facing conformation. *Proc. Natl. Acad. Sci. U. S. A.* 110, 8876–8881 (2013). doi:10.1073/pnas.1306849110
33. Smirnova, I., Kasho, V., Choe, J.Y., Altenbach, C., Hubbell, W.L., Kaback, H.R.: Sugar binding induces an outward facing conformation of LacY. *Proc Natl Acad Sci U S A.* 104, 16504–16509 (2007). doi:0708258104 [pii]r10.1073/pnas.0708258104
34. Kaback, H.R.: A chemiosmotic mechanism of symport. *Proc. Natl. Acad. Sci. U. S. A.* 112, 1259–64 (2015). doi:10.1073/pnas.1419325112
35. Jiang, X., Andersson, M., Chau, B.T., Wong, L.Y., Villafuerte, M.K.R., Kaback, H.R.: Role of Conserved Gly-Gly Pairs on the Periplasmic Side of LacY. *Biochemistry*. 55, 4326–4332 (2016). doi:10.1021/acs.biochem.6b00666
36. Jiang, X., Smirnova, I., Kasho, V., Wu, J., Hirata, K., Ke, M., Pardon, E., Steyaert, J., Yan, N., Kaback, H.R.: Crystal structure of a LacY–nanobody complex in a periplasmic-open conformation. *Proc. Natl. Acad. Sci.* 113, 12420–12425 (2016). doi:10.1073/pnas.1615414113
37. Holyoake, J., Sansom, M.S.P.: Conformational Change in an MFS Protein: MD Simulations of LacY. *Structure*. 15, 873–884 (2007). doi:10.1016/j.str.2007.06.004
38. Doerr, S., Giorgino, T., Martinez-Rosell, G., Damas, J.M., De Fabritiis, G.: High-throughput automated preparation and simulation of membrane proteins with HTMD. *J. Chem. Theory Comput.* acs.jctc.7b00480 (2017). doi:10.1021/acs.jctc.7b00480
39. Ash, W.L., Zlomislic, M.R., Oloo, E.O., Tieleman, D.P.: Computer simulations of membrane proteins, (2004)
40. Jensen, M.Ø., Yin, Y., Tajkhorshid, E., Schulten, K.: Sugar transport across lactose permease probed by steered molecular dynamics. *Biophys. J.* 93, 92–102 (2007). doi:10.1529/biophysj.107.103994
41. Mackerell laboratory, University of Maryland, Baltimore, USA
42. Departments of Biopharmaceutical Sciences and Pharmaceutical Chemistry, and California Institute for Quantitative Biomedical Research, Mission Bay Byers Hall, University of California San Francisco, San Francisco, CA 94143, USA.
43. University of Illinois at Urbana-Champaign, USA
44. Theoretical and computational biophysics group, USA
45. Jo, S., Kim, T., Iyer, V.G., Im, W.: CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* 29, 1859–1865 (2008). doi:10.1002/jcc.20945
46. University of Michigan, Ann Arbor, MI 48109 USA
47. Birkbeck, University of London, UK

48. Finkelstein, A. V., Ptitsyn, O.: Protein Physics: A Course of Lectures. (2002)
49. Alenghat, F.J., Golan, D.E.: Membrane protein dynamics and functional implications in mammalian cells. *Curr. Top. Membr.* 72, 89–120 (2013). doi:10.1016/B978-0-12-417027-8.00003-9
50. Madej, M.G.: Function, Structure, and Evolution of the Major Facilitator Superfamily: The LacY Manifesto. *Adv. Biol.* 2014, 1–20 (2014). doi:10.1155/2014/523591
51. Sargsyan, K., Grauffel, C., Lim, C.: How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* 13, 1518–1524 (2017). doi:10.1021/acs.jctc.7b00028
52. Findlay, H.E., Booth, P.J.: The folding, stability and function of lactose permease differ in their dependence on bilayer lipid composition. *Sci. Rep.* 7, (2017). doi:10.1038/s41598-017-13290-7
53. Serdiuk, T., Sugihara, J., Mari, S.A., Kaback, H.R., Müller, D.J.: Observing a lipid-dependent alteration in single lactose permeases. *Structure.* 23, 754–761 (2015). doi:10.1016/j.str.2015.02.009
54. Childers, M.C., Daggett, V.: Insights from molecular dynamics simulations for computational protein design. *Mol. Syst. Des. Eng.* 2, 9–33 (2017). doi:10.1039/C6ME00083E
55. Guan, L., Kaback, H.R.: Binding affinity of lactose permease is not altered by the H⁺ electrochemical gradient. *Proc. Natl. Acad. Sci. U. S. A.* 101, 12148–12152 (2004). doi:10.1073/pnas.0404936101
56. Park, S., Khalili-Araghi, F., Tajkhorshid, E., Schulten, K.: Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality. *J. Chem. Phys.* 119, 3559 (2003). doi:10.1063/1.1590311
57. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E.: The Protein Data Bank. *Nucleic Acids Res.* 28, 235–242 (2000)

Appendices

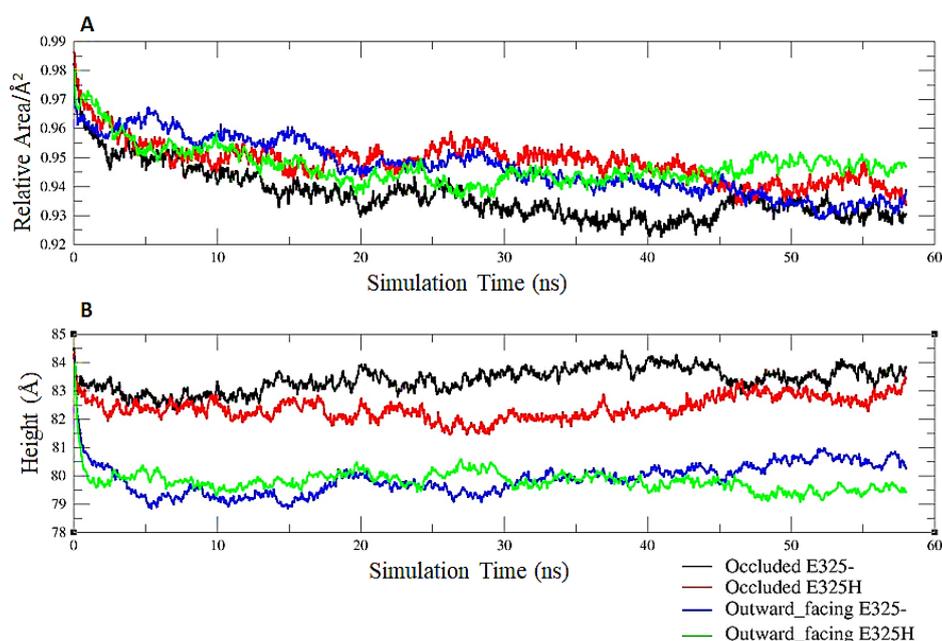


Fig 1. shows the equilibration of the simulated systems as referred from cell dimensions throughout the simulation time ~ 60 ns. The periodic cell parameters for the respective systems are shown in specific colours. Blue represents the outward-facing E325-, green represents the outward-facing E325H, red represents the occluded E325H and black represents the occluded E325- system. A represents the relative area/Å² of the simulation boxes for the respective systems and B represents the height/Å for the specific systems.

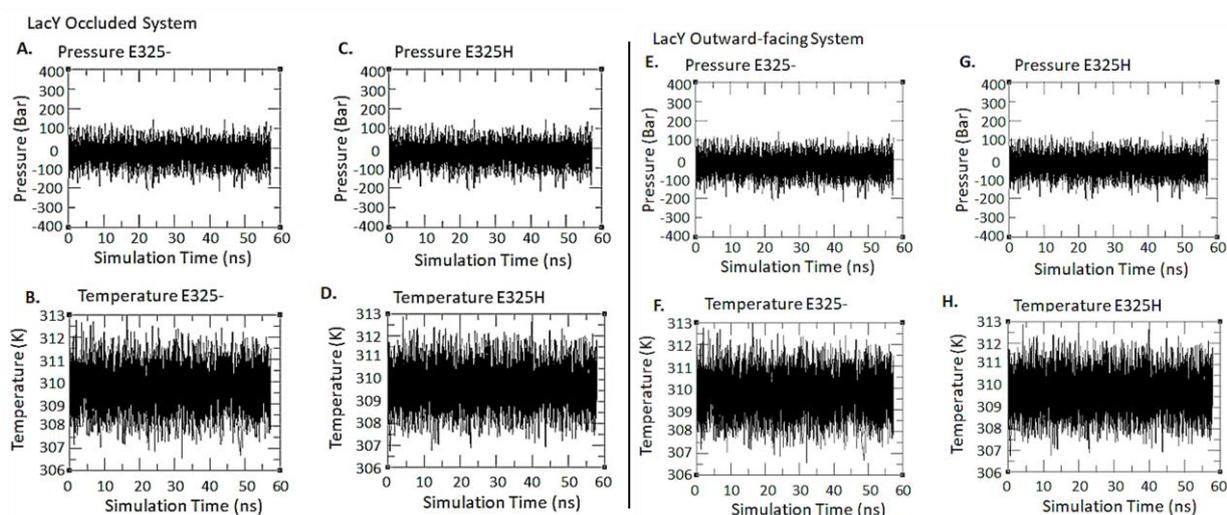


Fig 2. shows fluctuation and distribution of temperature and pressure of the LacY systems through the simulation time following the equilibration phase. The temperature in all systems (Fig 2B, 2D, 2G and 2H) is observed to fluctuate within the set value of 310 K. The pressure in the occluded system state (Fig 2A and 2C) fluctuates within the value of 1 Bar and 1 Bar for the outward-facing state system (Fig 2E and Fig 2F). The kinetics of the systems was observed to remain constant throughout the simulation time of ~ 60 ns.

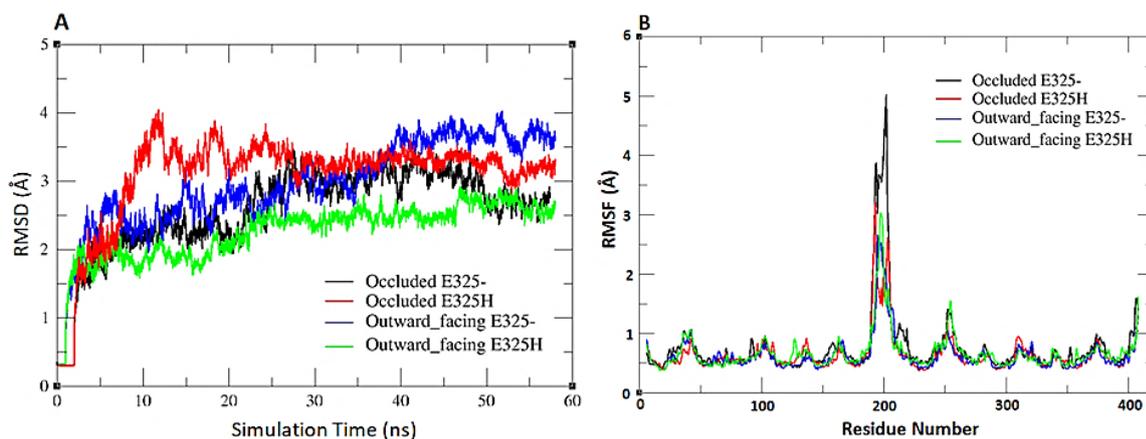


Fig 3. shows the rms distance with respect to the simulated LacY systems indicating their stability and equilibration. Figure 3A represents the root mean square deviation plots for backbone vs simulation time (~ 60 ns) for the four LacY systems, occluded E325- in black, occluded E325H in red, outward-facing E325- in blue and outward E325H in green. Figure 3B represents RMSF plots for backbone vs residue number for the four systems, occluded E325- in black, occluded E325H in red, outward-facing E325- in blue and outward E325H in green. The total number of residues considered was 417 amino acids for each system.

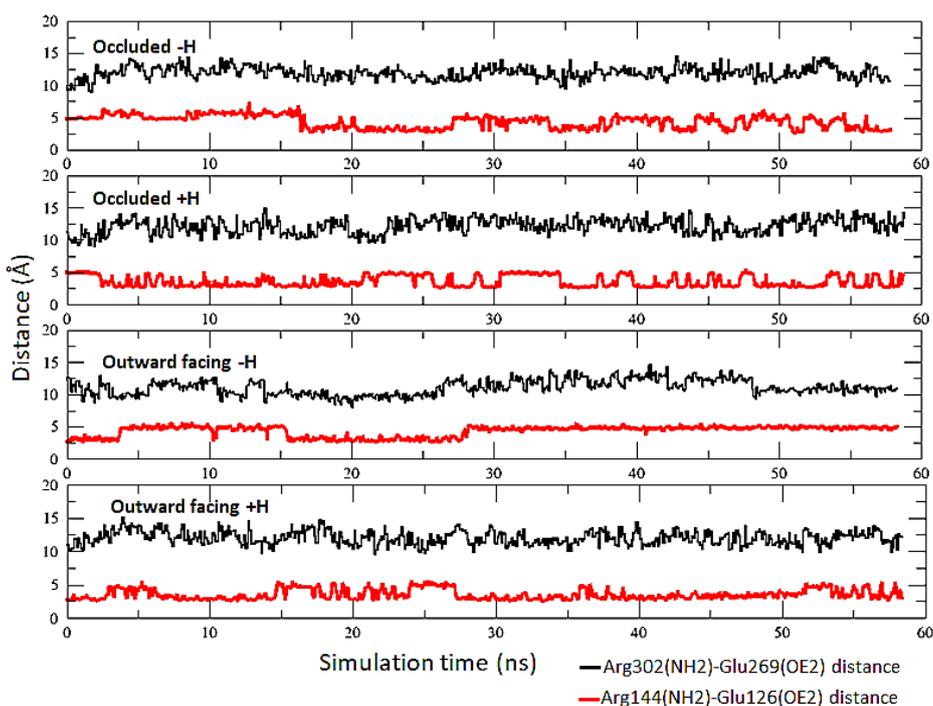


Fig 4. shows the inter atomic distances throughout the simulation time (~ 60 ns) for each simulated systems presented in the order of occluded E325- > occluded E325H > outward-facing E325- and outward-facing E325H. Black represents the hydrogen bond distance between Arg302(NH2)-Glu269(OE2) and red represents the hydrogen bond distance between Arg144(NH2)-Glu126(OE2).

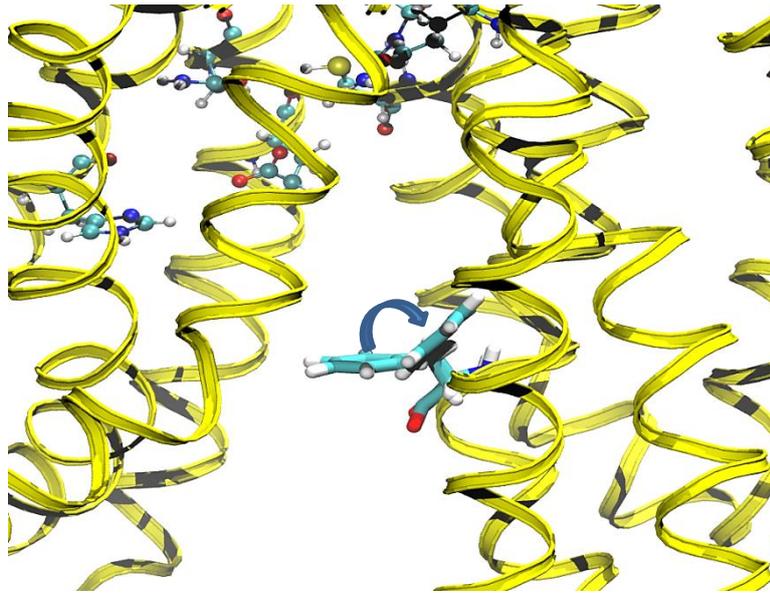


Fig 5. shows the cartoon presentation of the dihedral angle of Phe27 during sugar movement to the binding site. In yellow ribbons is the protein model structure of outward facing state, in cyan is the phe27 residue making an angle of $(-100$ to $100^\circ)$ and in balls and sticks are the residues in the binding pocket.

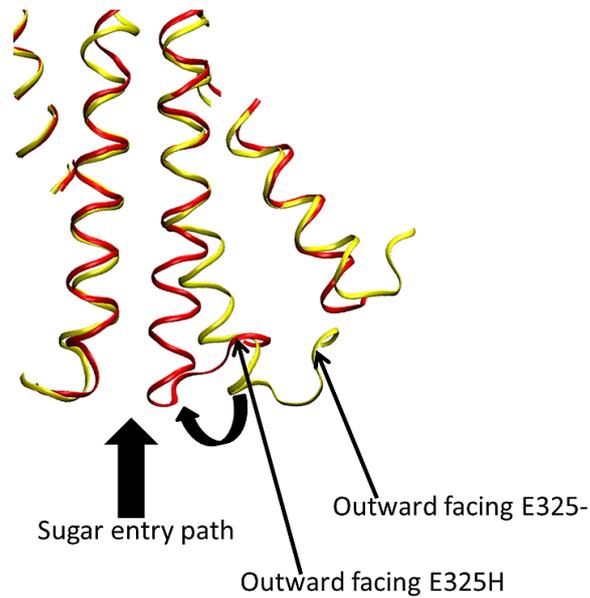


Fig 6. shows the helical shift (helix VII) in the outward-facing protonated system (E325H). Red represents the outward-facing system (E325H) and yellow represents the outward-facing system (E325-), the broad black arrow represents the entry pathway of the Lactose sugar.

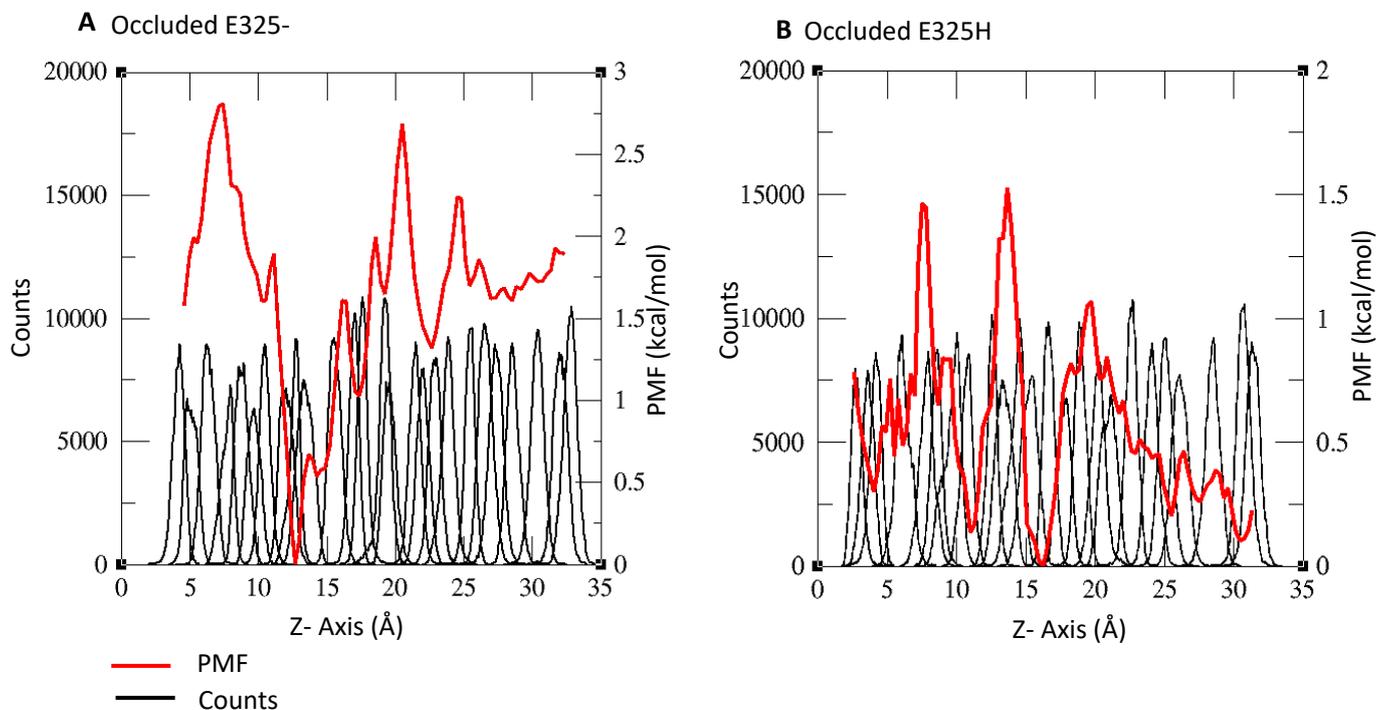


Fig 7. shows the potential of mean force for the occluded protonated and deprotonated systems. The red lines represent the PMF and the black line represents the counts in the simulated windows along the Z – Axis.

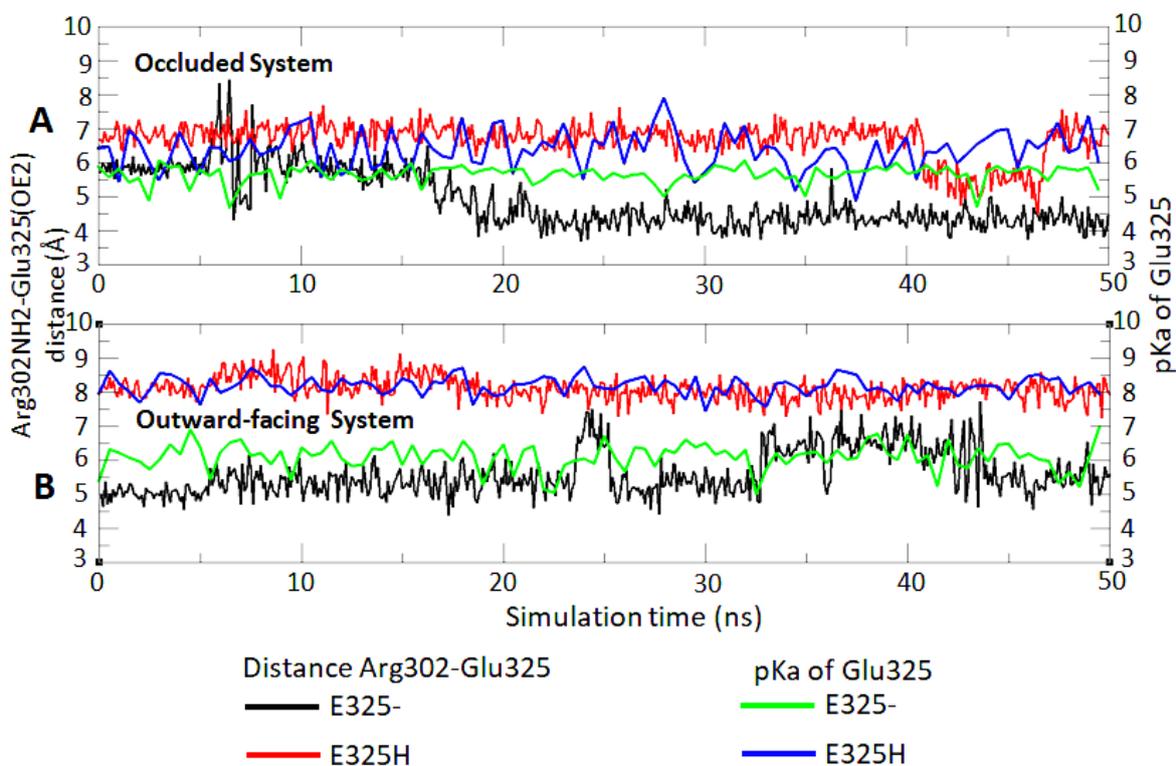


Fig 8. shows the interatomic distance between Arg302 and Glu325 and the pKa of Glu325. Fig. 8A represents the occlude system and Fig. 8B represents the outward facing system. The interatomic distance is shown in black for the deprotonated system (E325-) and in red for the protonated system (E3254H). The pKa of Glu325 is shown in green for the deprotonated system (E325-) and in blue for the protonated system (E325H).

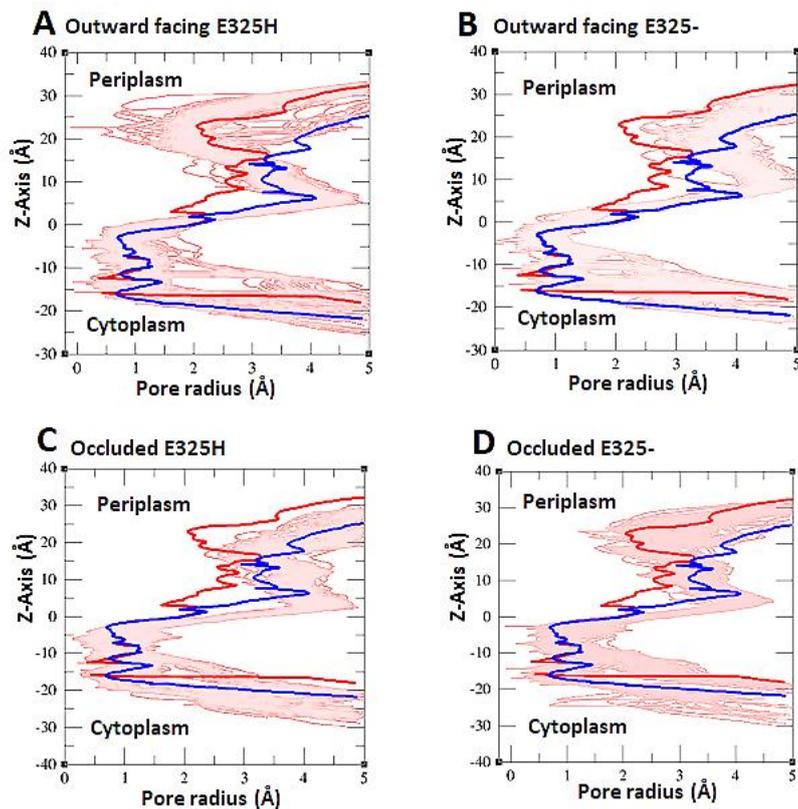


Fig 9. shows the variations in pore sizes for the four simulated LacY model systems with reference to the respective crystal structures PDBID: 4zyr (occluded conformation in thick red line) and PDBID 5GXB (outward-facing conformation thick blue line). In pink lines represents the water occupancy profiles of each system throughout the simulation time (10 ns). Fig. 9A shows the pore diameter for the outward facing E325H system, Fig. 9B for the outward facing E325- system, Fig. 9C represents the occluded E325H system and Fig. 9D representing the occluded E325- system. The periplasmic region is shown at the top and the cytoplasmic region is shown at the bottom of the diameter profile.